

122960

**STIC-Biotech/ChemLib**

---

**From:** Whiteman, Brian  
**Sent:** Wednesday, May 26, 2004 7:34 AM  
**To:** STIC-Biotech/ChemLib  
**Subject:** seq search

09/475,704 Barnett et al., 12/30/99

Please search SEQ ID NO: 1, 2, 3, and 4 against us patent and us patent publication databases.

Thank you,

Brian Whiteman  
Remsen, 2D14  
mail box 2C18  
Patent Examiner - Art Unit 1635  
United States Patent and Trademark Office  
(571) 272-0764

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: \_\_\_\_\_  
Date Completed: \_\_\_\_\_  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH:  
NA Sequences: \_\_\_\_\_  
AA Sequences: \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_



```
OY 241 ACCGTGCGCAACCTGTACTGCTGCAAGAGATCGAGGTCGCGGACACCAAGAGGCC 300
    |||||
DB 247 ACCGTGCGCAACCTGTACTGCTGCAAGAGATCGAGGTCGCGGACACCAAGAGGCC 306
OY 301 CTGGAACAAGATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
    |||||
DB 307 CTGGAAGAAGATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 366
OY 361 GCCG-----ACAAAGGCAAGGTGAGCCAGAACTACCCATGCTGCAAGAACTTG 408
    |||||
DB 367 GCCGCGGCAACCGCAACAGAGCCAGGTGAGCCAGAACTACCCATGCTGCAAGAACTTG 426
OY 409 CAGGCGCAGATGCTGCAACCAAGCCATCAGCCCGCACCTGAAAGCTGCTGGAAGGTG 468
    |||||
DB 427 CAGGCGCAGATGCTGCAACCAAGCCATCAGCCCGCACCTGAAAGCTGCTGGAAGGTG 486
OY 469 ATCGAGAGAGAGCCCTTCAAGCCCGAGGTGATCCCATGTTCAACGCGCTGAGAGGCG 528
    |||||
DB 487 GTGAGAGAGAGAGCCCTTCAAGCCCGAGGTGATCCCATGTTCAAGCGCTGAGAGGCG 546
OY 529 GCCACCCCGCAGAGACCTGAAACAAGTGTGAACACCGTGGCGGCACCAAGCGCCATG 588
    |||||
DB 547 GCCACCCCGCAGAGACCTGAAACAAGTGTGAACACCGTGGCGGCACCAAGCGCCATG 606
OY 589 CAGATGCTGAAGGACACCATCAACGAGAGGCGCGAGTGGACCGCGTGACCCCGTG 648
    |||||
DB 607 CAGATGCTGAAGGAGACCATCAACGAGAGGCGCGAGTGGACCGCGTGACCCCGTG 666
OY 649 CAGCGCGGCGCCATCGCCCGCGCAGATGCGGAGCGCCCGCGCAGCGACATCGCGGC 708
    |||||
DB 667 CAGCGCGGCGCCATCGCCCGCGCAGATGCGGAGCGCCCGCGCAGCGACATCGCGGC 726
OY 709 ACCACCAAGACCCCTGCAAGAGAGATGCTGATGACCAAGACCCCGCATCCCGTG 768
    |||||
DB 727 ACCACCAAGACCCCTGCAAGAGAGATGCTGATGACCAAGACCCCGCATCCCGTG 786
OY 769 GCGCAGATCTAACAAGCGGTGATCATCTGGGCTGAAACAAGATGCTGCGATGTACAG 828
    |||||
DB 787 GCGCAGATCTAACAAGCGGTGATCATCTGGGCTGAAACAAGATGCTGCGATGTACAG 846
OY 829 CCGGTGAGACATCTGACATCAAGAGAGGCGCCCAAGAGCCCTTCCGCACTACGTGAC 888
    |||||
DB 847 CCGGTGAGACATCTGACATCAAGAGAGGCGCCCAAGAGCCCTTCCGCACTACGTGAC 906
OY 889 CGCTTCTTCAAGACCTGCGCGCGAGAGAGAGACCCAGAGGTGAAGACTGATGACC 948
    |||||
DB 907 CGCTTCTTCAAGACCTGCGCGCGAGAGAGAGACCCAGAGGTGAAGACTGATGACC 966
OY 949 GACACCTTGTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1008
    |||||
DB 967 GAGACCTTGTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1026
OY 1009 CCGCGCGCAGCGCTGAGAGAGAGATGACCGCTGCAAGGCGGTGGCGGCCAGCGCAC 1068
    |||||
DB 1027 CCGCGCGCAGCGCTGAGAGAGAGATGACCGCTGCAAGGCGGTGGCGGCCAGCGCAC 1086
OY 1069 AAGCGCGCGGTGCTGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1122
    |||||
DB 1087 AAGCGCGCGGTGCTGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1146
OY 1123 CAGAAAGAGAACTTCAAGGCGCCCGCGCATGCTCAAGTGTCTTCACTGCGCAAGAG 1182
    |||||
DB 1147 CAGCGCGGCACTTCCGCAACCAAGGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1206
OY 1183 GCGCAATGCGCGCAACTGCGCGCGCCCGCGCAAGAGAGAGAGAGAGAGAGAGAG 1242
    |||||
DB 1207 GCGCAATGCGCGCAACTGCGCGCGCCCGCGCAAGAGAGAGAGAGAGAGAGAGAG 1266
OY 1243 GAGGCGCAACCAAGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1302
    |||||
DB 1267 GAGGCGCAACCAAGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1326
OY 1303 CCGAGCCACAAGGCGCGCGCGCACTTCTGCAAGCGCGCCGAGCCACCGCGCCC 1362
```

```
DB 1327 CCCAGCTACAAAGGCGCGCGCGCAACTTCTGCAAGAGCGCGCCGAGCCACCGCCCC 1386
OY 1363 CCGCGCAGAGACTTCCGCTTC-----GAGAGACCAACCCCGCGCAGAGAGAGAG-- 1413
DB 1387 CCGGAGAGAGACTTCCGCTTCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1446
OY 1414 ---AGCAAGAGCGGAGAGCCCTGACCAGCTGAAGAGCTGTTCGCAAGAGAGAGAGAG 1470
DB 1447 ATGACAAAGAGAGCTGTACCCCTGACCAAGCTGCGCAGCTGTTCGCAAGAGAGAG 1506
OY 1471 AGCCAGTAA 1479
DB 1507 AGCCAGTAA 1515

RESULT 2
US-09-475-515-75
; Sequence 75, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCES: 1621, 002
; CURRENT APPLICATION NUMBER: US/09/475, 515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 75
; LENGTH: 4472
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; US-09-475-515-75

Query Match      81.0%; Score 1197.8; DB 4; Length 4472;
Best Local Similarity 89.3%; Pred. No. 2.9e-173;
Matches 1347; Conservative 0; Mismatches 132; Indels 30; Gaps 4;

OY 1 ATGGGCGCGCGCGCAGCATCTGCGCGCGGCAAGCTGAGCGCTGGAGCGCATCCG 60
DB 2939 ATGGGCGCGCGCGCAGCGCTGAGCGCGGAGCTGGAACAAGTGGAGAAATCCG 2998
OY 61 CTGCGCGCGCGCGCAAGAGTGTACATGATGAAGCACTGTGTGGCCAGCGCGAG 120
DB 2999 CTGCGCGCGCGCGCAAGAGTGTACATGATGAAGCACTGTGTGGCCAGCGCGAG 3058
OY 121 CTGGAAGATTGCGCCCTGAACCCCGGCTGCTGAGAACAGCGAGGGCTGCAAGCATC 180
DB 3059 CTGGAAGCTTTCCTGTAACCCCGGCTGCTGAGAACAGCGAGGGCTGCGCGCATC 3118
OY 181 ATCGGCGAGCTGCAACCCCGCTGCAAGACCGGCGAGAGAGAGAGAGAGAGAGAG 240
DB 3119 CTGCGCGAGCTGAGCCAGCTGCAAGACCGGCGAGAGAGAGAGAGAGAGAGAGAG 3178
OY 241 ACCGTGCGCAACCTGTACTGCTGCAAGAGATCGAGGTCGCGGACACCAAGAGGCC 300
DB 3179 ACCGTGCGCAACCTGTACTGCTGCAAGAGATCGAGGTCGCGGACACCAAGAGGCC 3238
OY 301 CTGGAACAAGATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
DB 3239 CTGGAAGAAGATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3298
```





Db 3495 CAGGGCCAGATGTGTCAACGAGCCATCAAGCCCCGCAACCTTGAACGCTGTGAAGTGTG 3554  
QY 469 ATCGAGAGAGAGGCTTCAAGCCCCGAGGTATCCCATGTTCAAGGCTGAGGAGGC 528  
Db 3555 GTGAGAGAGAGGCTTCAAGCCCCGAGGTATCCCATGTTCAAGGCTGAGGAGGC 3614  
QY 529 GCCACCCCCCAGAGACTTGAACAGATGTTGAACACCTGGGCGGCAACGAGCCGCATG 588  
Db 3615 GCCACCCCCCAGAGACTTGAACAGATGTTGAACACCTGGGCGGCAACGAGCCGCATG 3674  
QY 589 CAGATGCTGAAGGACCACTCAACGAGAGGCGCGGAGTGGGACCGGTCACCCCGTG 648  
Db 3675 CAGATGCTGAAGGACCACTCAACGAGAGGCGCGGAGTGGGACCGGTCACCCCGTG 3734  
QY 649 CAGGCGGCGCCCATCGCCCCCGGCAAGTGGCGGAGCGCCCGGCAAGCATCGCGGC 708  
Db 3735 CAGGCGGCGCCCATCGCCCCCGGCAAGTGGCGGAGCGCCCGGCAAGCATCGCGGC 3794  
QY 709 ACCACCAAGCACTCTGCAAGAGCAAGATCGGCTGGATGACCAAGCAACCCCGCATCCCGTG 768  
Db 3795 ACCACCAAGCACTCTGCAAGAGCAAGATCGGCTGGATGACCAAGCAACCCCGCATCCCGTG 3854  
QY 769 GGGGACATCTAACAAGGGGTGATCATCTGTGGCTTGAACAAGATGTGTACAGC 828  
Db 3855 GGGGACATCTAACAAGGGGTGATCATCTGTGGCTTGAACAAGATGTGTACAGC 3914  
QY 829 CCGGTGAGATCTCTGACATCAAGAGGCCCCAAGAGGCTTCCGCACTACGTGAC 888  
Db 3915 CCGGTGAGATCTCTGACATCAAGAGGCCCCAAGAGGCTTCCGCACTACGTGAC 3974  
QY 889 CGCTTCTTCAAGACCTCTGCGCGCGGCAAGCAAGCAAGAGGTGAAGACTGTGATGAC 948  
Db 3975 CGCTTCTTCAAGACCTCTGCGCGCGGCAAGCAAGCAAGAGGTGAAGACTGTGATGAC 4034  
QY 949 GACACCTCTGCTGTGACAGAACGCCCAACCCCGCACTGCAAGCAATCTGTGCGCTCTCGGC 1008  
Db 4035 GACACCTCTGCTGTGACAGAACGCCCAACCCCGCACTGCAAGCAATCTGTGCGCTCTCGGC 4094  
QY 1009 CCGGCGCGCACTCTGAGAGAGATGATGACCGCTCTGCAAGGCGTGGGCGGCCAGCCAC 1068  
Db 4095 CCGGCGCGCACTCTGAGAGAGATGATGACCGCTCTGCAAGGCGTGGGCGGCCAGCCAC 4154  
QY 1069 AAGGCGCGCTGTGCGCGGAGGAGTGAAGCCAGGCCAACCA-----CCAGCTGATGATG 1122  
Db 4155 AAGGCGCGCTGTGCGCGGAGGAGTGAAGCCAGGCCAACCA-----CCAGCTGATGATG 4214  
QY 1123 CAGAAAGCACTTCAAGGGGCCCCCGGCGCATCTGCAAGTCTTCACTGCGGCAAGAG 1182  
Db 4215 CAGCGCGGCACTTCCGCAACCAAGCGAAGACCTTCAAGTCTTCACTGCGGCAAGAG 4274  
QY 1183 GGGCACAATGCGCGCAACTGCGCGGCGCCCCCGGCAAGAGGCGTCTGGAAGTGGCGCAAG 1242  
Db 4275 GGGCACAATGCGCGCAACTGCGCGGCGCCCCCGGCAAGAGGCGTCTGCGCGCGC 4334  
QY 1243 GAGGGCCACAGATGAAGGACTGCAACGAGCGCCAGGCCAACTTCTGGGCAAGATCTG 1302  
Db 4335 GAGGGCCACAGATGAAGGACTGCAACGAGCGCCAGGCCAACTTCTGGGCAAGATCTG 4394  
QY 1303 CCGAGCCACAAGGGCGCGCGCGGCAACTTCTGCAAGAGCGCGCCCGGCAAGCGCCCGC 1362  
Db 4395 CCGAGCTACAAGGGCGCGCGCGGCAACTTCTGCAAGAGCGCGCCCGGCAAGCGCCCGC 4454  
QY 1363 CCGCGCGAGAGCTTCCGCTTC-----GAGAGACCAACCCCGCGCAAGAGCAGAG-- 1413  
Db 4455 CCGGAGGAGAGCTTCCGCTTCCGCGGAGGAGAGACCAACCCCGCAAGAGCAGAGCCC 4514  
QY 1414 ---AGCAAGAGCGCGGAGACCTTGAACAGCCTGGAAGAGCCTGTTGGCAACGACCCCTG 1470  
Db 4515 ATGAGCAAGAGCTGTACCCCTTGAACAGCCTTGGCAGCCTGTTGGCAACGACCCCGAC 4574  
QY 1471 AGCCAGTAA 1479  
Db 4575 AGCCAGTAA 4583

RESULT 4  
US-09-475-515-74  
; Sequence 74, Application US/09475515A  
; Patent No. 6602705  
; GENERAL INFORMATION:  
; APPLICANT: BARNETT, Susan  
; APPLICANT: ZUR MEGEDE, Jan  
; APPLICANT: SRIVASTAVA, Indresh  
; APPLICANT: LIAN, Yang  
; APPLICANT: HARTOG, Karin  
; APPLICANT: GRBER, Catherine  
; APPLICANT: SELBY, Mark  
; APPLICANT: WALKER, Christopher  
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION  
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES  
; FILE REFERENCE: 1621.002  
; CURRENT APPLICATION NUMBER: US/09/475,515A  
; CURRENT FILING DATE: 1999-12-30  
; NUMBER OF SEQ ID NOS: 90  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 74  
; LENGTH: 4689  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:  
; OTHER INFORMATION: gp160.modsf162.gag.modsf2  
US-09-475-515-74  
Query Match 81.0%; Score 1197.8; DB 4; Length 4689;  
Best Local Similarity 89.3%; Pred. No. 2.9e-173;  
Matches 1347; Conservative 0; Mismatches 132; Indels 30; Gaps 4;  
QY 1 ATGGGCGCGCGCGCAGCATCTGCGCGCGCAAGCTGACGCTGGAGCGCATCCGC 60  
Db 3156 ATGGGCGCGCGCGCAGCATCTGCGCGCGCAAGCTGACGCTGGAGCGCATCCGC 3215  
QY 61 CTGCGCGCGCGCGCAAGAGTGTCTACATGATGAAGCACTGTGTGGCCAGCGCGAG 120  
Db 3216 CTGCGCGCGCGCGCAAGAGTGTCTACATGATGAAGCACTGTGTGGCCAGCGCGAG 3275  
QY 121 CTGAGAAATTCGCTTGAACCCCGCGCTGTGAGACCAAGAGGCTGCAAGATC 180  
Db 3276 CTGAGCGCTTTCGCTGTAACCCCGCGCTGTGAGACCAAGAGGCTGCGCCAGATC 3335  
QY 181 ATCCGCAAGTGCACCCCGCGCTGTGAGACCGGCAAGAGCTGAAGAGCTGTTCAAC 240  
Db 3336 CTGGGCGAGCTGACGCCAGCTGTGAGACCGGCAAGAGCTGCGAGCTGTACAAC 3395  
QY 241 ACCGTGGCACTCTGTACTGCGTGAACGAGATCGAGGTCCGCAACCAAGAGGCC 300  
Db 3396 ACCGTGGCACTCTGTACTGCGTGAACGAGATCGAGGTCCGCAACCAAGAGGCC 3455  
QY 301 CTGACAAGATGAGGAGGAGCAACAAGTGCAGAGCAAGATCCAGAGCGGAGGCC 360  
Db 3456 CTGAGAAAGATGAGGAGGAGCAACAAGTGCAGAGCAAGAGGCCAGAGGCCCGCC 3515  
QY 361 GCGG-----ACAAGGCAAGGTGAGCCAGAACTACCCCATCTGTGCAAGACTG 408  
Db 3516 GCGGCGGCAACCGGCAACAGAGCCAGGTGAGCCAGAACTACCCCATCTGTGCAAGACTG 3575  
QY 409 CAGGGCCAGATGTGCAACAGGCGCATCAGCCCCCGCAACCTGTAAGCCTGGTGAAGGTG 468  
Db 3576 CAGGGCCAGATGTGCAACAGGCGCATCAGCCCCCGCAACCTGTAAGCCTGGTGAAGGTG 3635  
QY 469 ATGAGGAGAGGCTTCAAGCGCGAGGTATCCCATGTTCAACCGCCTGAGAGGGC 528  
Db 3636 GTGAGAGAGAGGCTTCAAGCGCGAGGTATCCCATGTTCAACCGCCTGAGAGGGC 3695  
QY 529 GCCACCCCGCAGACCTGAACAGATGTTGAACACGTTGGCGGCCACCAAGCGCCCATG 588

```
|||||
Db 3696 GCCACCCCGGAGGACCTGAAACAGATGTTGAACACCGTGGCGGCCACCGAGCCGCATG 3755
Qy 589 CAGATGCTGAAGGACACCATCAACGAGAGCGCCGAGTGGGACCGGCTGACCCCGTG 648
Db 3756 CAGATGCTGAAGGAGACCATCAACGAGAGCGCCGAGTGGGACCGGCTGACCCCGTG 3815
Qy 649 CAGCGCGCGCCCATGCGCCCGCGGAGATGCGCGGAGCCCGCGGAGCGGACATCGCGGC 708
Db 3816 CAGCGCGCGCCCATGCGCCCGCGGAGATGCGCGGAGCCCGCGGAGCGGACATCGCGGC 3875
Qy 709 ACCACGAGACCCCTGAGAGAGATGCGCTGGATGACCAACCCCGCATCCCGCTG 768
Db 3876 ACCACGAGACCCCTGAGAGAGATGCGCTGGATGACCAACCCCGCATCCCGCTG 3935
Qy 769 GCGGACATCTAACAGCGGTGATCATCTGGGCTGAAACAGATCGTGCGATGTACAGC 828
Db 3936 GCGGAGATCTAACAGCGGTGATCATCTGGGCTGAAACAGATCGTGCGATGTACAGC 3995
Qy 829 CCGGTGAGATCTTGAACATCAAGGAGCGCCCAAGAGCGCCCTTCCGCGACTACGTGAC 888
Db 3996 CCGACGAGATCTTGAACATCTTGGCGCTGAGAGGCGCCCAAGAGCGCCCTTCCGCGACTACGTGAC 4055
Qy 889 CGCTTCTTCAAGACCTGCGCGCGCGGAGAGACACCCAGAGGTGAAGACTGATGATGAC 948
Db 4056 CGCTTCTTCAAGACCTTGGCGCTGAGAGGCGCCCAAGAGCGTGAAGACTGATGATGAC 4115
Qy 949 GACACCTTGTGTGTCAGAACGCGCAACCCCACTGCAAGACCATCTGCGCGCTCTGCGC 1008
Db 4116 GAGACCTTGTGTGTCAGAACGCGCAACCCCACTGCAAGACCATCTGAAAGCTCTGCGC 4175
Qy 1009 CCGGCGCGGAGCTTGAAGAGATGATGACCGGCTTGGCAAGGCGTGGGCGGCGCCAGCCAC 1068
Db 4176 CCGGCGCGGAGCTTGAAGAGATGATGACCGGCTTGGCAAGGCGTGGGCGGCGCCAGCCAC 4235
Qy 1069 AAGGCGCGGCTGTGGCGGAGCGATGAGCGGAGGCGCAACA-----CCAGCGTATGATG 1122
Db 4236 AAGGCGCGGCTGTGGCGGAGCGATGAGCGGAGGCGTGAAGCGGCGGAGCATCATGATG 4295
Qy 1123 CAGAAAGAGCACTTCAAGGCGCGCGCGCATCTGTAAGTCTTCAACTGCGGCAAGAG 1182
Db 4296 CAGCGCGGCACTTCCGCAACGAGCGGAAGACGCTCAAGTCTTCAACTGCGGCAAGAG 4355
Qy 1183 GCGCACATGCGCGGCAACTGCGCGCGCGCGCGCAAGAGGCGTGTGAAGTGGCGGCAAG 1242
Db 4356 GCGCACATGCGCGGCAACTGCGCGCGCGCGCGCGCAAGAGGCGTGTGAAGTGGCGGCGC 4415
Qy 1243 GAGGCGCACAGATGAAGGACTGCAACCGAGCGGCAAGGCGCACTTCTGGGCAAGATCTG 1302
Db 4416 GAGGCGCACAGATGAAGGACTGCAACCGAGCGGCAAGGCGCACTTCTGGGCAAGATCTG 4475
Qy 1303 CCGAGCGCAAGAGGCGCGCGCGGCAACTTCTGTCAGAGCGCGCGCGGAGCGGCAAGCGCC 1362
Db 4476 CCGAGCGCAAGAGGCGCGCGCGGCAACTTCTGTCAGAGCGCGCGCGGAGCGGCAAGCGCC 4535
Qy 1363 CCGCGCGGAGGCTTCCGCTT-----GAGGAGACCAACCCCGCGCGCAAGAGAGAG--- 1413
Db 4536 CCGGAGGAGGCTTCCGCTTCCGCGGAGGAGAGACCAACCCCGAGCGCAAGAGAGAGCGCC 4595
Qy 1414 ---AGCAAGGAGCGCGGAGACCTGACAGCGTGAAGAGCGTGTGGCAAGCAAGCGCCCTG 1470
Db 4596 ATGACAGAGAGCTGTACCTCCCTGACAGCTGCGCAGCTGTGTGGCAAGCAAGCGCCAGC 4655
Qy 1471 AGCCAGTAA 1479
Db 4656 AGCCAGTAA 4664
```

RESULT 5  
US-09-475-515-73  
; Sequence 73, Application US/09475515A  
; Patent No. 6602705  
; GENERAL INFORMATION:

```
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475, 515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 73
; LENGTH: 4766
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: gp160.modUs4.gag.modSf2
US-09-475-515-73
```

```
Query Match 81.0%; Score 1197.8; DB 4; Length 4766;
Best Local Similarity 89.3%; Pred. No. 2.9e-173;
Matches 1347; Conservative 0; Mismatches 132; Indels 30; Gaps 4;

Qy 1 ATGGGCGCGCGCGCGGAGAGAGTCTATCATGATGAAGCACTGTGTGGCGCGCGGAG 60
Db 3233 ATGGGCGCGCGCGCGGAGAGAGTCTATCATGATGAAGCACTGTGTGGCGCGCGGAG 3292
Qy 61 CTGGCGCGCGCGCGGAGAGAGTCTATCATGATGAAGCACTGTGTGGCGCGCGGAG 120
Db 3293 CTGGCGCGCGCGCGGAGAGAGTCTATCATGATGAAGCACTGTGTGGCGCGCGGAG 3352
Qy 121 CTGGAAGAGTTCGCGCTGTAACCCCGCGCTGTGAGAGACCGAGCGGCTGCAAGCATC 180
Db 3353 CTGGAAGAGTTCGCGCTGTAACCCCGCGCTGTGAGAGACCGAGCGGCTGCAAGCATC 3412
Qy 181 ATCCGCGAGCTGCAACCCCGCTGCAAGCGGCGAGCGGAGTGAAGAGCTGTTCAAC 240
Db 3413 CTGGGCGAGCTGCAAGCGGCGCTGCAAGCGGCGAGCGGAGTGAAGAGCTGTTCAAC 3472
Qy 241 ACCGTGCGCACCTGTATCTGCGTGCACGAGAGATGAGGTCGCGGAGCAAGAGGCG 300
Db 3473 ACCGTGCGCACCTGTATCTGCGTGCACGAGAGATGAGGTCGCGGAGCAAGAGGCG 3532
Qy 301 CTGACAAGATCGAGAGAGAGAGAGAGTCCAGAGAGATCCAGAGCGCGGAGCGG 360
Db 3533 CTGAGAAGATCGAGAGAGAGAGAGAGTCCAGAGAGAGAGAGAGCGCGCGGCGG 3592
Qy 361 GCGG-----ACAGGCGCAAGGTGAGCGCAAGCTACCTCATCTGTGAGAACTG 408
Db 3593 GCGGCGGCGCAACCGGCAAGCGAGCGAGGTGAGCGCAAGCTACCTCATCTGTGAGAACTG 3652
Qy 409 CAGGCGCAGATGTGTCACAGGCGCATCAGCGCGCGGAGCGCGCTGGAAGAGTG 468
Db 3653 CAGGCGCAGATGTGTCACAGGCGCATCAGCGCGCGGAGCGCGCTGGAAGAGTG 3712
Qy 469 ATGAGGAGAGGCGCTTCAAGCGCGGAGTGAATCCCATGTTCAAGCGCGCTGAGAGGCG 528
Db 3713 GTGAGGAGAGGCGCTTCAAGCGCGGAGTGAATCCCATGTTCAAGCGCGCTGAGAGGCG 3772
Qy 529 GCCACCGCGGAGAGCTGAAACAGATGTTGAACAGCGTGGCGGCGCACAGCGCGCATG 588
Db 3773 GCCACCGCGGAGAGCTGAAACAGATGTTGAACAGCGTGGCGGCGCACAGCGCGCATG 3832
Qy 589 CAGATGCTGAAGGACACCATCAACGAGAGCGCGCGAGTGGGACCGGCTGACCCCGTG 648
Db 3833 CAGATGCTGAAGGAGACCATCAACGAGAGCGCGCGAGTGGGACCGGCTGACCCCGTG 3892
```

```
QY 649 CAGCCGCGCCCATGCCCCCGCCAGATGCGGAGCCCCCGCGCAGCGACATCGCCGC 708
|||
DB 3893 CAGCCGCGCCCATGCCCCCGCCAGATGCGGAGCCCCCGCGCAGCGACATCGCCGC 3952
709 ACCACGAGCAACCTGAGAGAGAGATGCGCTGATGACCAAGACCCCGCCATCGCTG 768
|||
DB 3953 ACCACGAGCAACCTGAGAGAGAGATGCGCTGATGACCAAGACCCCGCCATCGCTG 4012
769 GCGGACATCTACAGCGGTGATCTGCGCTGAACAAGATGCTGCGATGACAGC 828
|||
DB 4013 GCGGACATCTACAGCGGTGATCTGCGCTGAACAAGATGCTGCGATGACAGC 4072
829 CCGGTGACATCTGACATCAAGAGGCGCCCAAGAGCCCTTCGCGACTACGTGAC 888
|||
DB 4073 CCGGTGACATCTGACATCAAGAGGCGCCCAAGAGCCCTTCGCGACTACGTGAC 4132
889 CCGTTCTTCAAGACCTTGGCGCGCGAGAGCAAGCAAGAGGTGAAGACTGATGACC 948
|||
DB 4133 CCGTTCTTCAAGACCTTGGCGCGCGAGAGCAAGCAAGAGGTGAAGACTGATGACC 4192
949 GACACCTGCTGTGACAGAACCGCAACCCCGACTGCAAGACCATCTGCGCTCTGCGC 1008
|||
DB 4193 GACACCTGCTGTGACAGAACCGCAACCCCGACTGCAAGACCATCTGCGCTCTGCGC 4252
1009 CCGCGCGCCAGCGCTGAGAGAGATGATGACCGCTGCAAGGCGGTGGCGCCCGAC 1068
|||
DB 4253 CCGCGCGCCAGCGCTGAGAGAGATGATGACCGCTGCAAGGCGGTGGCGCCCGAC 4312
1069 AAGGCGCGCTGCTGGCGCGAGGCGATGAGCCAGCCACA-----CCAGCTGATGATG 1122
|||
DB 4313 AAGGCGCGCTGCTGGCGCGAGGCGATGAGCCAGCCACA-----CCAGCTGATGATG 4372
1123 CAGAAGAGCACTTCAAGGCGCCCGCGCATGCTCAAGTGTCTCACTGCGGCAAGAG 1182
|||
DB 4373 CAGCGCGGCACTTCCGCAACCGAGGGAAGACCGTCAAGTGTCTCACTGCGGCAAGAG 4432
1183 GCGGCATGCGCCGCACTGCGCGCGCCCGCGCAAGAGGCTGCTGGAAGTGGCGCAAG 1242
|||
DB 4433 GCGGCATGCGCCGCACTGCGCGCGCCCGCGCAAGAGGCTGCTGGAAGTGGCGCAAG 4492
1243 GAGGGCCACAGATGAAGACTGACCGAGCGCGCAACTTCTGCGCAAGATCTG 1302
|||
DB 4493 GAGGGCCACAGATGAAGACTGACCGAGCGCGCAACTTCTGCGCAAGATCTG 4552
1303 CCGAGCCACAGAGGCGCGCGCGCAACTTCTGCGAGAGCGCGCGCAAGCGCGCC 1362
|||
DB 4553 CCGAGCTACAGAGGCGCGCGCGCAACTTCTGCGAGAGCGCGCGCAAGCGCGCC 4612
1363 CCGCGGAGAGCTTCCGCTT-----GAGGAGACCAAGCGCGCGCAAGAGAGAG-- 1413
|||
DB 4613 CCGGAGAGAGCTTCCGCTTCCGCGAGAGAGAGACCAAGCGCGCAAGAGAGAGCGCC 4672
1414 ---AGCAAGAGCGCGAGAGCGCTGACAGCTGAGAGCGCTGTTGCGCAAGAGCGCGCTG 1470
4673 ATGCAAGAGAGCTGATCCCGCTGACAGCGCTGCGAGCGCTGTTGCGCAAGAGCGCGCAGC 4732
QY 1471 AGCCAGTAA 1479
|||
DB 4733 AGCCAGTAA 4741
```

RESULT 6  
US-09-475-515-7  
Sequence 7, Application US/09475515A  
Patent No. 6602705

GENERAL INFORMATION:  
APPLICANT: BARNETT, Susan  
APPLICANT: ZUR MEGEDB, Jan  
APPLICANT: SRIVASTAVA, Indresh  
APPLICANT: LIAN, Ying  
APPLICANT: HARTOG, Karin  
APPLICANT: LIU, Hong  
APPLICANT: GRBER, Catherine

```
; APPLICANT: SELBY, Mark  
; APPLICANT: WALKER, Christopher  
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION  
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES  
; FILE REFERENCE: 1621.002  
; CURRENT APPLICATION NUMBER: US/09/475,515A  
; CURRENT FILING DATE: 1999-12-30  
; NUMBER OF SEQ ID NOS: 90  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 7  
; LENGTH: 2031  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: HIV-gag/HCV-core fusion polypeptide  
US-09-475-515-7
```

Query Match 80.9%; Score 1195.8; DB 4; Length 2031;  
Best Local Similarity 89.3%; Pred. No. 5.8e-173;  
Matches 1345; Conservative 0; Mismatches 132; Indels 30; Gaps 4;

```
QY 1 ATGGGCGCCCGCGCCAGCATCTGCGCGCGCGCAAGCTGAGCGCTGGAGCGCATCCG 60
|||
DB 7 ATGGGCGCCCGCGCCAGCATCTGAGCGCGCGCGCAAGCTGAGCGCTGGAGCGCATCCG 66
61 CTGCGCCCGCGCGCCAGAGAGTGTCTACATGATGAAGCACCTGTGTGGCCAGCGCGAG 120
|||
DB 67 CTGCGCCCGCGCGCCAGAGAGTGTCTACATGATGAAGCACCTGTGTGGCCAGCGCGAG 126
121 CTGAGAAAGTTCCTCTGAACCGCGCGCTGCTGAGAACCGAGCGAGCGCTGCAAGCATC 180
|||
DB 127 CTGAGAGCGCTTCCGCTGTAACCGCGCGCTGCTGAGAACCGAGCGAGCGCTGCGCATC 186
181 ATCCGCGAGCTGACCCCGCGCTGACAGACCGCGAGCGAGCGCTGAAAGCGCTGTTCAAC 240
187 CTGGGCGAGCTGACCGCGCTGACAGACCGCGAGCGAGCGCTGCGAGCGCTGTTCAAC 246
241 ACCGTGCGCACCTGTACTGCTGCTGACAGAGATGAGGTCCGCGCACCAAGAGGCGC 300
247 ACCGTGCGCACCTGTACTGCTGCTGACAGCGCATGACGTCAAGAGCACCAAGAGGCGC 306
QY 301 CTGAGCAAGATCGAGAGAGAGAGCAAGTGTCCAGAGAGATCCAGAGCGCGAGGCGC 360
|||
DB 307 CTGAGAAAGATCGAGAGAGAGAGCAAGTGTCCAGAGAGATCCAGAGCGCGCGCGC 366
361 GCGG-----ACAAGGCGAAGGTGAGCGAGAACTAACCCCATCTGTGAGAACCTG 408
|||
DB 367 GCGGCGCGCACCGGCAACAGACAGCGAGTGAAGCAAGTAACTAACCCCATCTGTGAGAACCTG 426
409 CAGGGCCAGATGTGACACCGCATCAGCGCGCGCACCTGTAAGCGCTGGTGAAGGTG 468
|||
DB 427 CAGGGCCAGATGTGACACCGCATCAGCGCGCGCACCTGTAAGCGCTGGTGAAGGTG 486
469 ATGAGGAGAGAGGCTTCAAGCGCGAGGTGATCCCATGTTCAAGCGCGCTGAGAGGCGC 528
487 GTGAGAGAGAGGCTTCAAGCGCGAGGTGATCCCATGTTCAAGCGCGCTGAGAGGCGC 546
529 GCCACCGCGCGAGAGCTGAAACAGATGTTGAACACCGTGGCGCGCACAGGCGCGCATG 588
547 GCCACCGCGCGAGAGCTGAAACAGATGTTGAACACCGTGGCGCGCACAGGCGCGCATG 606
QY 589 CAGATGCTGAAGAGCACCATCAACGAGAGGCGCGCGAGTGGAGCGCGCTGACCGCGTG 648
|||
DB 607 CAGATGCTGAAGAGACCATCAACGAGAGGCGCGCGAGTGGAGCGCGCTGACCGCGTG 666
649 CAGCGCGCGCGCATGCGCCCGCGCGAGATGCGCGAGCGCGCGCGAGCATGCGCGCGC 708
667 CAGCGCGCGCGCATGCGCCCGCGCGAGATGCGCGAGCGCGCGCGCGAGCATGCGCGCGC 726
QY 709 ACCACGAGCAACCTGAGAGAGAGATGCGCTGATGACCAAGCAAGCGCGCGCGCTG 768
|||
DB 727 ACCACGAGCAACCTGAGAGAGAGATGCGCTGATGACCAAGCAAGCGCGCGCGCTG 786
```

QY 769 GCGACATCTTCAAGCGGTGATCATCTGGGCTGAACAGATCGTGGATGTACAGC 828  
DB 787 GCGGAGATCTACAGCGGTGATCATCTGGGCTGAACAGATCGTGGATGTACAGC 846  
QY 829 CCCGTGAGCATCTTGAACATCAAGCAGGCGCCCAAGAGCCCTTCCGCACTACGTGAC 888  
DB 847 CCCACAGCATCTTGAACATCAAGCAGGCGCCCAAGAGCCCTTCCGCACTACGTGAC 906  
QY 889 CGCTTCTTCAAGACCTCTGCGCGCCGAGCAGACACCCAGAGGTGAAGACTGATGAC 948  
DB 907 CGCTTCTTCAAGACCTCTGCGCGCCGAGCAGACACCCAGAGGTGAAGACTGATGAC 966  
QY 949 GACACCTCTGCTGTGAGAGAGATGATGACCGCTGCGAGACCATCTGCGCTCTGCGC 1008  
DB 967 GAGACCTCTGCTGTGAGAGAGATGATGACCGCTGCGAGACCATCTGAGGCTCTGCGC 1026  
QY 1009 CCCGCGCCCACTTGAAGAGATGATGACCGCTGCGAGACCATCTGCGCTCTGCGC 1068  
DB 1027 CCCGCGCCCACTTGAAGAGATGATGACCGCTGCGAGACCATCTGAGGCTCTGCGC 1086  
QY 1069 AAGCGCGCTGTGCGCGAGATGATGACCGAGCCACACACACACACACACACACAC 1122  
DB 1087 AAGCGCGCTGTGCGCGAGATGATGACCGAGCCACACACACACACACACACACAC 1146  
QY 1123 CAGAGAGCATCTTCAAGGCGCCCGCGCATCTGCTCAAGTCTTCACTGCGCAAGAG 1182  
DB 1147 CAGCGCGCATCTTCAAGGCGCCCGCGCATCTGCTCAAGTCTTCACTGCGCAAGAG 1206  
QY 1183 GCGCACATCTGCGCGCATCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1242  
DB 1207 GCGCACATCTGCGCGCATCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1266  
QY 1243 GAGGCGCCACAGATGAAGAGCTGACCGAGCGCCAGAGCGCTGCTGAGAGTGGCG 1302  
DB 1267 GAGGCGCCACAGATGAAGAGCTGACCGAGCGCCAGAGCGCTGCTGAGAGTGGCG 1326  
QY 1303 CCCAGCCACAGAGCG 1362  
DB 1327 CCCAGCCACAGAGCG 1386  
QY 1363 CCCGCGCGCATCTTCCGCTTC-----GAGAGACCAACCCCGCGCGCGCGCGCG 1413  
DB 1387 CCCGCGCGCATCTTCCGCTTC-----GAGAGACCAACCCCGCGCGCGCGCGCG 1446  
QY 1414 ---AGCAGAGACCGCGAGACCTTACCAAGCTTGAAGAGCTTGGCAAGACCCCTG 1470  
DB 1447 ATGACAGAGAGCTGTACCCCTTACCAAGCTTGGCAAGCTTGGCAAGACCCCTG 1506  
QY 1471 AGCCAGT 1477  
DB 1507 AGCCAGT 1513

RESULT 7  
US-09-475-515-5  
; Sequence 5, Application US/09475515A  
; Patent No. 6602705  
; GENERAL INFORMATION:  
; APPLICANT: BARNETT, Susan  
; APPLICANT: ZUR MEGDE, Jan  
; APPLICANT: SRIVASTAVA, Indresh  
; APPLICANT: LIAN, Ying  
; APPLICANT: HARTOG, Karin  
; APPLICANT: LIU, Hong  
; APPLICANT: GREER, Catherine  
; APPLICANT: SBLBY, Mark  
; APPLICANT: WALKER, Christopher  
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION  
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES  
; FILE REFERENCE: 1621.002  
; CURRENT APPLICATION NUMBER: US/09/475,515A  
; CURRENT FILING DATE: 1999-12-30

NUMBER OF SEQ ID NOS: 90  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 1853  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; US-09-475-515-5

Query Match 75.4%; Score 1114.6; DB 4; Length 1853;  
Best Local Similarity 85.8%; Pred. No. 1.1e-160;  
Matches 1295; Conservative 0; Mismatches 184; Indels 30; Gaps 4;

QY 1 ATGGCGCGCGCGCGCGCATCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60  
DB 7 ATGGCGCGCGCGCGCGCGCATCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 66  
QY 61 CTGCG 120  
DB 67 CTGCG 126  
QY 121 CTGAGAGATCTGCG 180  
DB 127 CTGAGAGATCTGCG 186  
QY 181 ATCCG 240  
DB 187 CTGCG 246  
QY 241 ACCGTGCG 300  
DB 247 ACCGTGCG 306  
QY 301 CTGAGAGATCTGAG 360  
DB 307 CTGAGAGAGATCTGAG 366  
QY 361 GCGG-----ACAGAGCGAGGTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 408  
DB 367 GCGGCG 426  
QY 409 CAGGCGCGAGATGTGCAACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 468  
DB 427 CAGGCGCGAGATGTGCAACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 486  
QY 469 ATCGAGAGAGAGCGCTTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 528  
DB 487 GTGAGAGAGAGCGCTTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 546  
QY 529 GCCACCG 588  
DB 547 GCCACCG 606  
QY 589 CAGATCTTGAAG 648  
DB 607 CAGATCTTGAAG 666  
QY 649 CAGCG 708  
DB 667 CAGCG 726  
QY 709 ACCACCGAGAGCGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 768  
DB 727 ACCACCGAGAGCGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 786  
QY 769 GCGGAGATCTTCAAGCGCGGTGATCATCTGGGCTGAACAGATCGTGGATGTACAGC 828  
DB 787 GCGGAGATCTTCAAGCGCGGTGATCATCTGGGCTGAACAGATCGTGGATGTACAGC 846  
QY 829 CCCGTGAGCATCTTGAACATCAAGCAGGCGCCCAAGAGCCCTTCCGCACTACGTGAC 888



```
Db 847 CCCACGACATCTGACATCCGCGCAGGGCCCAAGAGGCCCTTCGCGACTACGTGAC 906
QY 889 CGCTTCTTCAAGACCCCTGCGCGCCGAGCAGACACCAGAGGTGAAGAACTGATGACC 948
Db 907 CGCTTCTTCAAGACCCCTGCGCGCTGAGCAGGCGCAGCAGACGTGAAGAACTGATGACC 966
QY 949 GACACCTGCTGTGAGAGAGAGATGATGACCGCTGCGCAGAGCCATCTGCGCTCTGCGC 1008
Db 967 GAGACCTGCTGTGAGAGAGAGATGATGACCGCTGCGCAGAGCCATCTGAGAGCTCTGCGC 1026
QY 1009 CCGCGCGCAGCCTGAGAGAGATGATGACCGCTGCGCAGAGCGCTGCGCGCGCCAGCCAC 1068
Db 1027 CCGCGCGCAGCCTGAGAGAGATGATGACCGCTGCGCAGAGCGCTGCGCGCGCCAGCCAC 1086
QY 1069 AAGCCCGCGCTGCTGCGCAGAGCGATGAGCCAGCCCAACA-----CCAGCGTATGATG 1122
Db 1087 AAGCCCGCGCTGCTGCGCAGAGCGATGAGCCAGCCCAACA-----CCAGCGTATGATG 1146
QY 1123 CAGAAAGACCACTTCAAGGCGCCCGCGCGCATCTGCAAGTGTCTTCAACTGCGCAAGAG 1182
Db 1147 CAGCGCGGCACTTCCGCAACACGCGGAAAGACGCTCAAGTGTCTTCAACTGCGCAAGAG 1206
QY 1183 GGCACCATGCGCGCAACTGCGCGCGCCCGCGCAAGAGGCTGCGCAAGTGTGCGCAAG 1242
Db 1207 GGCACCATGCGCGCAACTGCGCGCGCCCGCGCAAGAGGCTGCGCGCGCGCGCGCGCGC 1266
QY 1243 GAGGCGCACCAAGATGAGAGACTGACCGAGCGCGCAAGGCTGCGCAAGTGTGCGCAAG 1302
Db 1267 GAGGCGCACCAAGATGAGAGACTGACCGAGCGCGCAAGGCTGCGCAAGTGTGCGCAAG 1326
QY 1303 CCCAGCCCAAGAGGCGCGCGCGCGCAACTTCTGCAAGCGCGCGCGCGCGCGCGCGCGC 1362
Db 1327 CCTTCTTCAAGGAGAGCGCGCGCGCGCAACTTCTTCAAGCGAGCGCGCGCGCGCGCGCA 1386
QY 1363 CCGCGCGAGAGCTTCCGCTTCAAGAG-----ACCAACCGCGCGCGCGCGCGCGCGCGC 1413
Db 1387 CCAGAGAGAGCTTCAAGTGTGCGGAGAGAAACAACCTCTCTCAAGAGCAGAGCGCGC 1446
QY 1414 ---AGCAAGAGCGCGAGACCTGACCGAGCGCGTGAAGAGCGTGTGCGCAAGAGCGCGCTG 1470
Db 1447 ATAGACAGAGACTGTATCTTAACTTCTCTCAGATCACTCTTGGCAAGAGCGCGCTCG 1506
QY 1471 AGCCAGTAA 1479
Db 1507 TCACAGTAA 1515
```

```
RESULT 8
US-09-475-515-78
; Sequence 78, Application US/09475515A
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAM, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LITV, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SIBLEY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 78
; LENGTH: 1865
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
```

OTHER INFORMATION: Description of Artificial Sequence: GPI  
US-09-475-515-78

```
Query Match 75.4%; Score 1114.6; DB 4; Length 1865;
Best Local Similarity 85.8%; Pred. No. 1.1e-160;
Matches 1295; Conservative 0; Mismatches 184; Indels 30; Gaps 4;

QY 1 ATGGGCGCGCGCGCAGCATCTGCGCGCGCGCAAGCTGAGCGCTGGAGCGCATCCGC 60
Db 13 ATGGGCGCGCGCGCGCAGCGCTGAGCGCGCGCGCGCAAGCTGAGCGCGCGCGCGCGCGC 72
QY 61 CTGCGCGCGCGCGCGCAAGAGTGTCTATGATGAGCACTGTGTGTGGCGCAGCGCGAG 120
Db 73 CTGCGCGCGCGCGCGCAAGAGTGTCTATGATGAGCACTGTGTGTGGCGCAGCGCGAG 132
QY 121 CTGGAAGATTGCGCTGAAACCGCGCGCTGCTGAGACCGAGCGGCTGCAAGCATC 180
Db 133 CTGGAAGCTTGTGCGGTGAACCGCGCGCTGCTGAGACCGAGCGGCTGCGCGCATC 192
QY 181 ATCCGCGAGCTGCAACCGCGCGCTGAGACCGCGCGAGAGCTGAAGCGCTGTCAAC 240
Db 193 CTGCGCGCGCGCGCGCAAGAGTGTCTATGATGAGCACTGTGTGTGGCGCAGCGCGAG 252
QY 241 ACCGTGGCGCACCTCTGTACTGCTGTGACGAGAAATCGAGTCCGCGCAACCAAGAGCGC 300
Db 253 ACCGTGGCGCACCTCTGTACTGCTGTGACGAGAAATCGAGTCCGCGCAACCAAGAGCGC 312
QY 301 CTGGAAGATTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
Db 313 CTGGAAGATTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 372
QY 361 GCGC-----ACAAGGCAAGGTGAGCGCAAGTACCCCATCTGTGAGAACTG 408
Db 373 GCGCGCGCGCACCGCGCAACAGCAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 432
QY 409 CAGGCGCAGATGTGCAACAGCGCATCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 468
Db 433 CAGGCGCAGATGTGCAACAGCGCATCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 492
QY 469 ATCGAGAGAGAGCGCTTCAAGCGCGCGAGTATCCCATGTTCAACCGCGCTGAGCGAGCGC 528
Db 493 GTGAGAGAGAGAGCGCTTCAAGCGCGCGAGTATCCCATGTTCAACCGCGCTGAGCGAGCGC 552
QY 529 GCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 588
Db 553 GCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 612
QY 589 CAGATGCTGAAGAGCAACCATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 648
Db 613 CAGATGCTGAAGAGAGCAACCATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 672
QY 649 CAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 708
Db 673 CAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 732
QY 709 ACCACCAAGACCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 768
Db 733 ACCACCAAGACCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 792
QY 769 GCGCAATCTCAAGCGGTGATCATCTGCGCGCTGAAACAAGATCTGTGGGATGTACAGC 828
Db 793 GCGCAATCTCAAGCGGTGATCATCTGCGCGCTGAAACAAGATCTGTGGGATGTACAGC 852
QY 829 CCGGTGAGCATCTGAGACATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 888
Db 853 CCGGTGAGCATCTGAGACATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 912
QY 889 CGCTTCTTCAAGACCTGCGCGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 948
Db 913 CGCTTCTTCAAGACCTGCGCGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 972
QY 949 GACACCTGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1008
```

Db	973	GAGACCCTGCTGCTGTGCAGAACGCCAACCCCGCACTGCAGAACCATCTGAAAGCTCTCGGC	10322
Qy	1009	CCCGCGCCAGCCTGAGAGAGATGATGACCGCCTGCAGAGGCGTGGCGGCCCCAGCCAC	10688
Db	1033	CCCGCGCCACCTCTGAGAGAGATGATGACCGCCTGCAGAGGCGTGGCGGCCCCAGCCAC	10922
Qy	1069	AAGGCCCGCTGCTGCGCCGAGGCGATGAGCCAGGCCAACAAAA-----CCAGCTGATGATG	11222
Db	1093	AAGGCCCGCTGCTGCGCCGAGGCGATGAGCCAGGCCAACAAAA-----CCAGCTGATGATG	11522
Qy	1123	CAGAAAGCAACTTCAGAGGCGCCCGCGCATGTCAGAGTCTTCAACTGCGGCAAGGAG	11822
Db	1153	CAGCGCGCAACTTCGCGAACCCAGCGGAAGACCGTCAGAGTCTTCAACTGCGGCAAGGAG	12122
Qy	1183	GGCCACATCGCCCGCAACTGCGCGCGCCCGCGCAAGAGGCGCTGCTGGAAGTGGCGAAG	12422
Db	1213	GGCCACACCGCCAGAACTGCGCGCGCCCGCGCAAGAGGCGCTGCTGCGCGCGCGC	12722
Qy	1243	GAGGGCCACCAATGAAAGGACTGCAACCGAGCGCCAGGCCAACTTCTGCGCAAGATCTGG	13022
Db	1273	GAGGACACCAATGAAAGATGCACTGAGAGACAGGCTAAATTTTAAAGGAAAGATCTGG	13322
Qy	1303	CCGAGCCACAAGGGCGCGCCCGGCAACTTCTGCAAGAGCGCGCTCCGAGGCCACCGCCCCC	13622
Db	1333	CCTTCTACAAAGGGAAGGCGCAGGGAATTTTCTTCAGAGCAGACCAAGAGCCAAAGACCCCA	13922
Qy	1363	CCCGCCGAGAGCTTCCGCTTGCAGAG-----ACCAACCCCGGCGCAGAAAGCAGAG--	14132
Db	1393	CCAGAAAGAGCTTCAGGTTTGGGGAGGAGAAAACAATCCCTCTCAAGAGCAGAGCGCG	14522
Qy	1414	---AGCAAGGACCGCGAGACCCTGACCAAGCCTGAAAGAGCTGTTGGCAAGCAACCCCTG	14702
Db	1453	ATAGACAAAGAACTGTATCTTTAACTTCCCTCAGATCACTCTTGGCAAGCAACCCCTG	15122
Qy	1471	AGCCAGTAA 1479	
Db	1513	TCACAGTAA 1521	

```

RESULT 9
US-09-475-515-79
; Sequence 79, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; PLS REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 79
; LENGTH: 1865
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
;
; OTHER INFORMATION: Description of Artificial Sequence: GP2
;
US-09-475-515-79

```

Query Match	75.4%	Score 1114.6;	DB 4;	Length 1865;
Best Local Similarity	85.8%;	Pred. No. 11e-160;		
Matches 1295; Conservative	0;	Mismatches 184;	Indels 30;	Gaps 4;
by	1	ATGGGCGCCCGCCAGCATCTTCGCGCGCGCAAGCTGGACCGCTGGAGCGCATCCGC	60	

D	b	13	ATGGGCGCCCGCGCAGCGTGTGAGCGGCGGAGCTGACAAAGTGGAGAAATCCGC	72
Q	y	61	CTGGCCCCCGCGCAAGAGTCTACATGATGAAGCACCTGGTGTGGCCAGCCGCGAG	120
D	b	73	CTGGCCCCCGCGCGCAAGAAAGTACAAAGCTGAAGCACATCTGTGTGGCCAGCCGCGAG	132
Q	y	121	CTGGAGAAGTTCCCTCCTGAACCCCGGCTGTGAGACCAAGGAGGGCTGCAAGCATTC	180
D	b	133	CTGGAGCGCTTCGCTGTGAACCCCGGCTGTGAGACCAAGGAGGGCTGCGCCAGATTC	192
Q	y	181	ATCCGCCAGCTGACCCCGCCTGACAGCCGCGAGGAGCTGAAGAGCTTTCAAC	240
D	b	193	CTGGGCCAGCTGACGCCAGCTGACAGCCGCGAGGAGGCTGTGCGCAGCTGTACAAC	252
Q	y	241	ACCGTGGCCACCCCTGTACTGTCGTGCACGAGAAGATCGAGTCCCGACACCAAGGAGGCC	300
D	b	253	ACCGTGGCCACCCCTGTACTGTCGTGCACGAGCGCATCGAGTCAAGGACACCAAGGAGGCC	312
Q	y	301	CTGGACAAGATCGAGGAGGAGCAACAAGTCCAGCAGAAGATCCAGAGGCCGAGGCC	360
D	b	313	CTGGAGAAGATCGAGGAGGAGCAACAAGTCCAGAGAAGGCCCAGAGGCCGCGCC	372
Q	y	361	GCCG-----ACAAGGCAAGGTGAGCCAGAACTACCTGCTGCAACCTG	408
D	b	373	GCCGCGCGGCAACCGGCAACAGCAGCCAGGTGAGCCAGAACTACCTGCTGCAACCTG	432
Q	y	409	CAGGGCCAGATGGTGCACCAGGSCCATCAGCCCCCGCACCTGAAAGCCTGGGTGAAGGTG	468
D	b	433	CAGGGCCAGATGGTGCACCAGGSCCATCAGCCCCCGCACCTGAAAGCCTGGGTGAAGGTG	492
Q	y	469	ATCGAGGAGAAAGCCTTCAAGCCCCGAGGTGATCCCATGTTCAACGCGCTGAGCGAGGC	528
D	b	493	GTGGAGGAGAAAGCCTTCAAGCCCCGAGGTGATCCCATGTTCAAGCGCTGAGCGAGGC	552
Q	y	529	GCCACCCCCCAAGACTGAAACAAGTGTGAACACGTTGGGGGCCAACAGGCGCCCATG	588
D	b	553	GCCACCCCCCAAGACTGAAACAAGTGTGAACACGTTGGGGGCCAACAGGCGCCCATG	612
Q	y	589	CAGATGCTGAAGGACACCATCAACGAGAGGCCCGCGAGTGGGACCGGTCACCCCGTG	648
D	b	613	CAGATGCTGAAGGAGACCATCAACGAGAGGCCCGCGAGTGGGACCGGTCACCCCGTG	672
Q	y	649	CAGCGCGGCCCCCATGCCCCCGGCGCAGATGGCGGAGCCCCCGGCGCAGGCATCGCCCGGC	708
D	b	673	CAGCGCGGCCCCCATGCCCCCGGCGCAGATGGCGGAGCCCCCGGCGCAGGCATCGCCCGGC	732
Q	y	709	ACCAACCAAGACCTTGCAAGGACAGATCGCTGGATGACCAAGAACCCCCCATCCCCCGTG	768
D	b	733	ACCAACCAAGACCTTGCAAGGAGCAGATCGGCTGGATGACCAACAACCCCCCATCCCCCGTG	792
Q	y	769	GCGGACATCTACAAGCGGTGATCATCTGGGCTGAAACAAGATCTGTGGATGTACAGC	828
D	b	793	GCGGAGATCTACAAGCGGTGATCATCTGGGCTGAAACAAGATCTGTGGATGTACAGC	852
Q	y	829	CCCGTGAGCATCTTGACATCAAGCAGGGCCCCAAGGAGCCCTTCGCGACTAAGTGAAC	888
D	b	853	CCCAACCAAGCATCTTGACATCTCGGCAAGGGCCCCAAGGAGCCCTTCGCGACTAAGTGAAC	912
Q	y	889	CGCTTCTTCAAGACCCCTGCGCGCGCAGCAGACACCAAGAGGTGAAGACTGATGACC	948
D	b	913	CGCTTCTTCAAGACCCCTGCGCGCTGAGCAGGCCAGCCAAGACTGAAGACTGATGACC	972
Q	y	949	GACACCCCTGCTGTGAGAACGCCAACCCGACTGCAAGACCATCTGCGCGCTTCGCGC	1008
D	b	973	GAGACCCCTGCTGTGAGAACGCCAACCCGACTGCAAGACCATCTGAAGGCTTCGCGC	1032
Q	y	1009	CCCGGCGCAGCCTGAGAGAGATGATGACCGCTGCAAGGCGTGGCGGCCCCAGGCCAC	1068
D	b	1033	CCCGGCGCAGCCTGAGAGAGATGATGACCGCTGCAAGGCGTGGCGGCCCCAGGCCAC	1092
Q	y	1069	AAGGCGCGCTGTGGCCAGGCGATGAGCCAGGCCACA-----CCAGCTGATGATG	1122



Db 1093 AAGCCCGCTGCTGCGGAGCGGATGAGCCAGTACCAACCCGGCCATCATGATG 1152  
Qy 1123 CAGAAAGCAACTTCAAGGCCCCCGCGCATGCTCAAGTGTCTTAAGTGGCAAGAG 1182  
Db 1153 CAGCGCGCACTTCCGCAACGAGCGGAGACCGTCAAGTGTCTTAAGTGGCAAGAG 1212  
Qy 1183 GGCCACATGCGCCGCAACTGCGCGCGCGCGCGGCAAGAGGCTGTGGAAGTGGCAAG 1242  
Db 1213 GGCCACATGCGCGGAGACTGCGCGCGCGCGCGGCAAGAGGCTGTGCGCGCGCG 1272  
Qy 1243 GAGGCGCAAGATGAGAGCTGCAACGAGCGCGCGCAACTTCTGCGCAAGATGCG 1302  
Db 1273 GAGGCGCAAGATGAGAGCTGCAACGAGCGCGCGCAACTTCTGCGCAAGATGCG 1332  
Qy 1303 CCCAGCGCAAGGCGCGCGCGCGCAACTTCTGCGCAAGCGCGCGCGCGCGCGCG 1362  
Db 1333 CTTCTTCAAGGAGGAGCGGAGGAGTCTTCAAGCGAGACGAGCGCAAGCGCGCG 1392  
Qy 1363 CCCGCGGAGCTTCCGCTTCAAGAG-----ACCAACCGCGCGCGCGCGCGCG 1413  
Db 1393 CCAGAGAGAGCTTCAAGTCTTGGGAGAGAAACAACCTCTCTCAAGAGAGAGCG 1452  
Qy 1414 ---AGCAGAGCGCGAGACCTGACAGCTGAGAGCTGTTCGCAAGAGCGCGCG 1470  
Db 1453 ATAGACAGAGACTGTATCTTAACTTCCCTCAGATCACTCTTGGCAAGAGCGCG 1512  
Qy 1471 AGCCAGTAA 1479  
Db 1513 TCACAGTAA 1521

RESULT 10  
US-09-475-515-6  
; Sequence 6, Application US/09475515A  
; Patent No. 6602705  
; GENERAL INFORMATION:  
; APPLICANT: BARNETT, Susan  
; APPLICANT: ZUR MEGEDE, Jan  
; APPLICANT: SRIVASTAVA, Indresh  
; APPLICANT: LIAN, Ying  
; APPLICANT: HARTOG, Karin  
; APPLICANT: LIU, Hong  
; APPLICANT: GREER, Catherine  
; APPLICANT: SELBY, Mark  
; APPLICANT: WALKER, Christopher  
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION  
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES  
; FILE REFERENCE: 1621.002  
; CURRENT APPLICATION NUMBER: US/09/475,515A  
; CURRENT FILING DATE: 1999-12-30  
; NUMBER OF SEQ ID NOS: 90  
; SOFTWARE: Patent Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 4319  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURES:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: HIV-Gag-polymerase  
US-09-475-515-6

Query Match 75.4%; Score 1114.6; DB 4; Length 4319;  
Best Local Similarity 85.8%; Pred. No. 1.1e-160;  
Matches 1295; Conservative 0; Mismatches 184; Indels 30; Gaps 4;

Qy 1 ATGGGCGCGCGCGCAGCATCTGCGCGCGCGGAGCTGAGCGCTGGAGCGCATCCG 60  
Db 7 ATGGGCGCGCGCGCAGCATCTGCGCGCGCGGAGCTGAGCAAGTGGAGAGATCCG 66  
Qy 61 CTGCGCGCGCGCGCAAGAGTCTACATGATGAGCACTGTGTGGCGCGCGAG 120  
Db 67 CTGCGCGCGCGCGCAAGAGTCTACATGATGAGCACTGTGTGGCGCGCGAG 126

Qy 121 CTGAGAGATTGCGCTGAAACCCCGCTGTGAGACGAGGAGGCTCAAGCAGATC 180  
Db 127 CTGAGCGCTTGGCGGTGAACCCCGCTGTGAGACGAGGAGGCTGCGCAGATC 186  
Qy 181 ATCCGCGACTGACCCCGCTGCAAGACCGGCGAGGAGCTGAAGAGCTGTTCAAC 240  
Db 187 CTGCGCGACTGACCGCGCTGCAAGACCGGCGAGGAGCTGCGCAGCTGTACAA 246  
Qy 241 ACCGTGCGCACTGTATCTGCTGCAACGAGAGATCGAGTCCGCGCAACGAGAG 300  
Db 247 ACCGTGCGCACTGTATCTGCTGCAACGAGAGATCGAGTCCGCGCAACGAGAG 306  
Qy 301 CTGAGAGATGAG 360  
Db 307 CTGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 366  
Qy 361 GCGG-----ACAGGCGCAAGTGAAGCGCAAGTGAAGTGAAGTGAAGTGA 408  
Db 367 GCGGCGCGCACTGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 426  
Qy 409 CAGGCGCAGATGTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 468  
Db 427 CAGGCGCAGATGTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 486  
Qy 469 ATGAG 528  
Db 487 GTGAG 546  
Qy 529 GCCACCGCGCAG 588  
Db 547 GCCACCGCGCAG 606  
Qy 589 CAGATGCTGAAG 648  
Db 607 CAGATGCTGAAG 666  
Qy 649 CAGCG 708  
Db 667 CAGCG 726  
Qy 709 ACCACAG 768  
Db 727 ACCACAG 786  
Qy 769 GCGGAGATCTAAG 828  
Db 787 GCGGAGATCTAAG 846  
Qy 829 CCGGTGAG 888  
Db 847 CCGGTGAG 906  
Qy 889 GCGTCTTCAAG 948  
Db 907 GCGTCTTCAAG 966  
Qy 949 GAGACCTGTGTGAG 1008  
Db 967 GAGACCTGTGTGAG 1026  
Qy 1009 CCGGCGCGCAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1068  
Db 1027 CCGGCGCGCAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1086  
Qy 1069 AAGCG 1122  
Db 1087 AAGCG 1146  
Qy 1123 CAG 1182  
Db 1147 CAGCGCGCGCACTTCCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1206  
Qy 1183 GCGCAGATGCGCGCGCAAGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 1242

Db	1207	GGCCACACCGCCAGAGAACTGCGCCGCCGCCCGCCAGAGAGGGCTGCGCGCTGCGGCCG	1266
Qy	1243	GAGGGCCACCCAGATGAAGGACTGCACCCGAGCGCCAGGCCAACCTTCTGGGCAAGATCTGG	1302
Db	1267	GAAAGACACCAATGAAGATTGCATGAGAGACAGGCTAATTTTTTAGGGAGATCTGG	1326
Qy	1303	CCGAGCCACAAGGGGCGCCCCCGGCACTTCTGCAAGAGCGCGCCGAGCCACCGCCCC	1362
Db	1327	CCTTCTTACAAGGGGAAGGCCAGGGAAATTTTCTTCAGAGCAGACCAGAGCCACAGCCCCA	1386
Qy	1363	CCCGCCGAGAGCTTTCGCTTCGAGGAG-----ACCACCCCGGGCCAGAAAGCAGGAG---	1413
Db	1387	CCAGAAAGAGAGCTTCAGGTTTGGGGAGGAGAAACAACCTCCCTCTCAGAAAGCAGGAGCGG	1446
Qy	1414	---AGCAAGGACCGGAGAGCCCTGACCAGCCTGAAGAGCCTGTTCGGCAACGACCCCTTG	1470
Db	1447	ATAGACAAAGGAATGTATCTTTAACTTCCCTCAGATCACTCTTTGGCAACGACCCCTCG	1506
Qy	1471	AGCCAGTAA	1479
Db	1507	TCACAGTAA	1515

## RESULT 11

```

US-09-475-515-9
; Sequence 9, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 1268
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic Gag
; OTHER INFORMATION: common region
US-09-475-515-9

```

Query Match 69.3%; Score 1025.2; DB 4; Length 1268;

Best Local Similarity	89.6%	Pred. No. 3.5e-147;
Matches 1131; Conservative	0;	Mismatches 113; Indels 18; Gaps 2;

Oy 1 ATGGCGCCCGCCAGCATCTTCGCGCGCGCAAGCTGACCGCTTGAGCGCATCCG 60  
 Db 7 ATGGCGCCCGCCAGCGTGTGAGCGCGCGGCGAGTGGACAAGTGGAGAAATCCG 66  
 Oy 61 CTGCGCCCCCGCGCAAGAGTGTACATGATGAAGCACTGTGTGCGCCAGCCGCGAG 120  
 Db 67 CTGCGCCCCCGCGCAAGAGAGTACAAGCTGAAGCACATGTGTGCGCCAGCCGCGAG 126  
 Oy 121 CTGAGAGAGTTCGCCCCCTGAACCCCGGCTGTGAGACCAGCGAGGCTGCAAGCAGATC 180  
 Db 127 CTGAGAGCGTTTCGCGCTGAACCCCGGCTGTGAGACCAGCGAGGCTGCGCCAGATC 186  
 Oy 181 ATCCGCCAGCTGCACCCCGCCTGCGAGACCGGCAGCGAGAGCTGAAGAGCTGTTCAAC 240  
 Db 187 CTGGGCAGCTGCAGCCCAAGCCTTGACAGACCGGCAGCGAGGAGCTGCGCAGCTGTACAC 246

QY	241	ACCGTGCCACCCTGTACTGCGTGCACGAGAGATCGAGGTCGCGACACCAAGAGGCC	300
Db	247	ACCGTGCCACCCTGTACTGCGTGCACGAGCATTCGACTCAAGACACCAAGAGGCC	306
QY	301	CTGGAACAAGATCGAGAGAGAGAGCAAACTGTCAGAGAGATTCAGCAGGCCGAGGCC	360
Db	307	CTGAGAGAGATCGAGAGAGAGAGCAAACTGTCAGAGAGAGAGAGAGAGAGAGAGAG	366
QY	361	GCCG-----ACAAAGGCAAGTGAAGCCAGAACTACCCATCGTGCAAACTG	408
Db	367	GCCGCGCGCACCGGCAACAGACAGCCAGGTGAACAGAACTACCCATCGTGCAAACTG	426
QY	409	CAGGCGCAGATGCTGCACCAAGCCATCAGCCCCCGCACCTGAAACGCTGCTGAAAGTG	468
Db	427	CAGGCGCAGATGCTGCACCAAGCCATCAGCCCCCGCACCTGAAACGCTGCTGAAAGTG	486
QY	469	ATCGAGAGAGAGGCTTTCAGCCCCCGAGGTGATCCCATGTTTCAACGCGCTGAGCGAGGC	528
Db	487	GTCGAGAGAGAGGCTTTCAGCCCCCGAGGTGATCCCATGTTTCAACGCGCTGAGCGAGGC	546
QY	529	GCCACCCCCCGAGAGACTGAAACAGATGTTTAAACACCGTGGCGGCCAACAGGCCGCATG	588
Db	547	GCCACCCCCCGAGAGACTGAAACAGATGTTTAAACACCGTGGCGGCCAACAGGCCGCATG	606
QY	589	CAGATGCTGAAGAGACACCATCAACGAGAGAGCCCGGAGTGGAGACCGCTGCAACCCCGTG	648
Db	607	CAGATGCTGAAGAGAGACCATCAACGAGAGAGCCCGGAGTGGAGACCGCTGCAACCCCGTG	666
QY	649	CACGCGGCGCCCATCGCCCCCGGCGAGATGCGCGAGAGCCCGCGGCGAGCGACATCGCCGCG	708
Db	667	CACGCGGCGCCCATCGCCCCCGGCGAGATGCGCGAGAGCCCGCGGCGAGCGACATCGCCGCG	726
QY	709	ACCACCAAGACCCCTGCAGAGCAGATCGCTGCTGATGACCAACACCCCCCATCCCGTG	768
Db	727	ACCACCAAGACCCCTGCAGAGCAGATCGCTGCTGATGACCAACACCCCCCATCCCGTG	786
QY	769	GCGCAGATCTTCAAGCGCGTGGATCATCTCGGCGCTGAACAAGATCGTGGCGATGTACAGC	828
Db	787	GCGCAGATCTTCAAGCGCGTGGATCATCTCGGCGCTGAACAAGATCGTGGCGATGTACAGC	846
QY	829	CCCGTGAGCATTCCTGACATCAAGCAGAGGCCCAAGGAGCCCTTCGCGACTACGTGAC	888
Db	847	CCCAACCAAGCATTCCTGACATCAAGCAGAGGCCCAAGGAGCCCTTCGCGACTACGTGAC	906
QY	889	CGCTTCTTCAAGACCTTGCGCGCGCGAGCAGACACCCAGAGGTGAAGAACTGATGACC	948
Db	907	CGCTTCTTCAAGACCTTGCGCGCGCGAGCAGACACCCAGAGGTGAAGAACTGATGACC	966
QY	949	GACACCCCTGCTGTGCAGAACGCCAACCCCGACTGCAAGAACCATCTGCGCGCTCTGCGC	1008
Db	967	GAGACCCCTGCTGTGCAGAACGCCAACCCCGACTGCAAGAACCATCTGCAAGAGCTCTGCGC	1026
QY	1009	CCCGGCGCGCAGCTGAGAGAGATGATGACCGCTGCCAGGCGCTGGCGCGCCAGCCAC	1068
Db	1027	CCCGGCGCGCAGCTGAGAGAGATGATGACCGCTGCCAGGCGCTGGCGCGCCAGCCAC	1086
QY	1069	AAGGCCCCGCTGTGCGCGAGGCGATGAGCCAGGCCAACCA-----CCAGCTGATGATG	1122
Db	1087	AAGGCCCCGCTGTGCGCGAGGCGATGAGCCAGGTGACGAACCCGCGCATCATGATG	1146
QY	1123	CAGAGAGCAACTTCAAGGAGCCCCCGCGCATGCTCAAGTCTTCAACTGCGGCAAGAG	1182
Db	1147	CAGCGCGCAACTTCCGCAACAGCGGAAGACCGTCAAGTCTTCAACTGCGGCAAGAG	1206
QY	1183	GCGCAGATGCGCGCAACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1242
Db	1207	GCGCAGATGCGCGCAACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1266
QY	1243	GA 1244	
Db	1267	GA 1268	

QY	1243	GA	1244
Db	1267	GA	1268

RESULT 12  
US-09-552-950-2  
; Sequence 2, Application US/09552950  
; Patent No. 6541248  
; GENERAL INFORMATION:  
; APPLICANT: Oxford Biomedica (UK) Limited  
; TITLE OF INVENTION: Anti-Viral Vectors  
; FILE REFERENCE: 674524-2004  
; CURRENT APPLICATION NUMBER: US/09/552,950  
; CURRENT FILING DATE: 2000-04-20  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 4307  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:gapool-synbp - codon  
; OTHER INFORMATION: optimised gapool sequence  
US-09-552-950-2

Query Match 64.1%; Score 947.6; DB 4; Length 4307;  
Best Local Similarity 80.7%; Pred. No. 1.8e-135;  
Matches 1135; Conservative 0; Mismatches 259; Indels 12; Gaps 2;

```
QY 1 ATGGGCGCCGCGCCAGCATCTGCGCGCGCAAGCTGAGCGCTGGAGCGATCCGC 60
   |||||
DB 1 ATGGGCGCCGCGCCAGCATCTGCGCGCGCAAGCTGAGCGCTGGAGAGATCCGC 60

QY 61 CTGCGCCCGCGCGCAAGAGTCTACATGATGAAGCACTGTGTGGCGCAGCCGAG 120
   |||||
DB 61 CTGCGCCCGCGCGCAAGAGTCTACATGATGAAGCACTGTGTGGCGCAGCCGAG 120

QY 121 CTGAGAGAGTTCGCGCTGAACTCCGCGCTGCTGAGAGACCAAGAGGCTGCAAGCATC 180
   |||||
DB 121 CTGAGAGAGTTCGCGCTGAACTCCGCGCTGCTGAGAGACCAAGAGGCTGCAAGCATC 180

QY 181 ATCCGCGAGCTGCAACCCCGCTGCAAGCCGCGAGCGAGAGCTGAAGAGCTGTCAAC 240
   |||||
DB 181 ATCCGCGAGCTGCAACCCCGCTGCAAGCCGCGAGCGAGAGCTGAAGAGCTGTCAAC 240

QY 181 CTGCGCAATGCAAGCCAGCTGCAAAACCGGAGCGAGAGCTGCGCAGCTGTACAAC 240
   |||||
DB 181 CTGCGCAATGCAAGCCAGCTGCAAAACCGGAGCGAGAGCTGCGCAGCTGTACAAC 240

QY 241 ACCGTGCGCACTCTGTACTGCTGCAAGAGATCGAGGTCCGACCAAGAGAGGCC 300
   |||||
DB 241 ACCGTGCGCACTCTGTACTGCTGCAAGAGATCGAGGTCCGACCAAGAGAGAGGCC 300

QY 241 ACCGTGCGCACTCTGTACTGCTGCAAGAGATCGAGGTCCGACCAAGAGAGAGGCC 300
   |||||
DB 241 ACCGTGCGCACTCTGTACTGCTGCAAGAGATCGAGGTCCGACCAAGAGAGAGGCC 300

QY 301 CTGGAACAAGATCGAGAGAGAGAGAGAGAGAGAGAGAGATCCAGAGAGCGGAGCC 360
   |||||
DB 301 CTGGAACAAGATCGAGAGAGAGAGAGAGAGAGAGAGAGATCCAGAGAGCGGAGCC 360

QY 301 CTGGAATAAATCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
   |||||
DB 301 CTGGAATAAATCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360

QY 361 GCC-----GACAAAGGCAAGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 414
   |||||
DB 361 GACACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420

QY 415 CAGATGTGCAACAGGCGCATAGCCCGCGAGCTGAAACGCTGGGTGAAGGTATCGAG 474
   |||||
DB 421 CAGATGTGCAACAGGCGCATAGCCCGCGAGCTGAAACGCTGGGTGAAGGTATCGAG 480

QY 475 GAGAGAGCTTCAGCCCGAGGTGATCCCATGTTCAACCGCTGAGCGAGAGCGCCACC 534
   |||||
DB 481 GAGAGAGCTTCAGCCCGAGGTGATCCCATGTTCTCAGCCCTGTCAGAGGAGCCACC 540

QY 535 CCCGAGGACCTGAACAGATGTTGAACACCGTGGGGCGGACAGAGCGGCGCATGAGATG 594
   |||||
DB 541 CCCGAGGACCTGAACAGATGTTGAACACAGTGGGGGAGACAGAGCGGCGCATGAGATG 600

QY 595 CTGAAGAGACCATCAACAGAGAGCGCGCGAGTGGAGACCGCGTGCACCCCGTGCAGCGC 654
   |||||
DB 601 CTGAAGAGACCATCAATGAGAGAGCGTGCAGATGGAGATCGTGTGCATCCGGTGCAGCGCA 660

QY 655 GGGCCCATGCGCCCGCGCAAGTGGCGAGAGCGCGCGGAGAGAGATGCGCGGACACACC 714
   |||||
DB 661 GGGCCCATGCGCGCGCGCAAGTGGCGTGAAGCGAGCGGGGCTCAGACATGCGCGGAGAGACT 720
```

```
QY 715 AGCACCTTGAGAGAGAGATGCTCTGATGACCAAGCAACCCCATATCCCGTGGCGGAC 774
   |||||
DB 721 AGTACCTTGAGAGAGAGATGCTCTGATGACCAAGCAACCCCATATCCCGTGGAGAGAA 780

QY 775 ATCTACAAAGCGGTGATCATCTGGGCTGGAACAAGATGTCGGATGTACAGCCCGTG 834
   |||||
DB 781 ATCTACAAAGCGGTGATCATCTGGGCTGGAACAAGATGTCGGATGTATAGCCCTTACC 840

QY 835 AGCATCTGACATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 894
   |||||
DB 841 AGCATCTGACATCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900

QY 895 TTCAAGACCTTGGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 954
   |||||
DB 901 TACAAAACGCTCCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960

QY 955 CTGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1014
   |||||
DB 961 CTGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020

QY 1015 GCCAGCTGAGAGAGAGATGATGACCGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAG 1074
   |||||
DB 1021 GCTACCTGAGAGAGAGATGATGACCGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080

QY 1075 CGCGTCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1128
   |||||
DB 1081 CGCGTCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140

QY 1129 AGCAACTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1188
   |||||
DB 1141 GCGAAGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200

QY 1189 ATGCGCCGCACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1248
   |||||
DB 1201 ACAGCCCGCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260

QY 1249 CACGAGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1308
   |||||
DB 1261 CACGAGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320

QY 1309 CACAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1368
   |||||
DB 1321 TACAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380

QY 1369 GAGAGCTTCGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1394
   |||||
DB 1381 GCGAGCTTCAAGTCTGCGGTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1406
```

RESULT 13  
US-09-552-950-5  
; Sequence 5, Application US/09552950  
; Patent No. 6541248  
; GENERAL INFORMATION:  
; APPLICANT: Oxford Biomedica (UK) Limited  
; TITLE OF INVENTION: Anti-Viral Vectors  
; FILE REFERENCE: 674524-2004  
; CURRENT APPLICATION NUMBER: US/09/552,950  
; CURRENT FILING DATE: 2000-04-20  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 9772  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:psynbp  
US-09-552-950-5

Query Match 63.5%; Score 939.8; DB 4; Length 9772;  
Best Local Similarity 78.6%; Pred. No. 2.7e-134;  
Matches 1182; Conservative 0; Mismatches 297; Indels 24; Gaps 4;

```
QY 1 ATGGGCGCCGCGCAGCATCTGCGCGGCGCAAGCTGACGCTGGGAGCGCATCCGC 60
DB 1108 ATGGGCGCCGCGCAGCGTGTGCGGGGCGAGCTGACCGCTGGGAGAAGATCCGC 1167
QY 61 CTGCGCCCGCGCGCAAGTGTCTACATGATGAAGACCTGTGTGGGCGAGCGCGAG 120
DB 1168 CTGCGCCCGCGCGCAAAAGATGACAGCTGAAGACATCGTGTGGGCGAGCGGAA 1227
QY 121 CTGGAAGATTGCGCTGAACCCCGGCTGTGAGACGAGCGAGGCTGAAGAGCTTTCAAC 180
DB 1228 CTGGAAGCTTGTGCGGTGAACCCCGGCTGTGAGACGAGCGAGGCTGCGAGCTGTACAC 1287
QY 181 ATCCGCAAGCTGACCCCGCTGCAAGACCGGCGAGGAGCTGAAGAGCTTTCAAC 240
DB 1288 CTGCGGCACTGACGCGCGCTGCAACCGGCGAGCGAGGAGCTGCGAGCTGTACAC 1347
QY 241 ACCGTGCGCAACCTGTACTGCGGTGCAAGAGATGAGTCCGCGCAACAAGAGGCC 300
DB 1348 ACCGTGCGCAACCTGTACTGCGGTGCAAGAGATGAGTCCGCGCAACAAGAGGCC 1407
QY 301 CTGCAAGAATCGAGAGGAGAGAGAGAGAGAGAGATCCAGAGGCC----- 354
DB 1408 CTGGAATAAATCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1467
QY 355 GAGGCGCGCGCAAGGCGAGGTGAGCGCAAGACTACCCCATCTGTGCAAGACTGAGGCG 414
DB 1468 GACACCGGACACAGCAACAGGTGAGCGCAAGACTACCCCATCTGTGCAAGACTGAGGCG 1527
QY 415 CAGATGTGCAACGAGGCTATCAGCCCGCGCAACCTGAACCTGTGTAAGGTGATGAG 474
DB 1528 CAGATGTGCAACGAGGCTATCAGCCCGCGCAACCTGTGTAAGGTGATGAG 1587
QY 475 GAGAGGCTTTCAGCGCGCGAGGTGATCCCATGTTCAACCGCTGTGAGCGCGCGCAC 534
DB 1588 GAGAGGCTTTCAGCGCGCGAGGTGATCCCATGTTCTCAGCCCTGTGAGAGGAGCGCAC 1647
QY 535 CCCCAGAGACTGAACAGCATGTTGAACACCGTGGGCGCGCACCGCGCATGAGATG 594
DB 1648 CCCCAGAGACTGAACAGCATGTTGAACACAGTGGGCGCGCACCGCGCATGAGATG 1707
QY 595 CTGAAGACACCATCAACAGAGAGGCGCGCGAGTGGGACCGCGCTGCAACCCCTGCAAGCC 654
DB 1708 CTGAAGAGACCATCAATGAGAGAGGCTCGGATGGGATGCTGTGATCCGCTGCAAGCG 1767
QY 655 GCGCCCATCGCCCCCGCGCGAGATGCGCGAGCGCGCGAGCATCGCGCGCACCGCAC 714
DB 1768 GCGCCCATCGCACCGCGCGAGATGCGGTAGCGCGCGCGCTCAAGACATCGCGCGAGACT 1827
QY 715 AGCACCTTCAGAGAGAGAGATGCTGTGATGACCAAGCAACCCCATCCCGTGGCGGAC 774
DB 1828 AGTACCTTCAGAGAGAGATGCTGTGATGACCAAGCAACCCCATCCCGTGGGAGAA 1887
QY 775 ATCTAAGACGCGGTGATCATCTGGGCTTGAACAGATGTCGGATGTACAGCCCGCTG 834
DB 1888 ATCTAAGACGCTGATCATCTGGGCTTGAACAGATGTCGGATGTACAGCCCTACC 1947
QY 835 AGCATCTTGAACATCAAGAGAGGCGCGCGAGAGCGCTTCCGCGACTACGTGAGACCGCTTC 894
DB 1948 AGCATCTTGAACATCCGCGCAAGGCGCGCGAGAGAGCGCTTCCGCGACTACGTGAGACCGGTTTC 2007
QY 895 TTCAAGACCTTCGCGCGCGCGAGAGAGCAACCAAGAGGTGAAGAACTGTGATGACCGACACC 954
DB 2008 TACAAAACGCTCCGCGCGAGAGGCTAGCGAGAGGTGAAGAACTGTGATGACCGAAACC 2067
QY 955 CTGCTGTGTCAGAAAGCGCAACCCCGAGCTGCAAGACCATCTGCGCGCTCTGCGCCCGGC 1014
DB 2068 CTGCTGTGTCAGAAAGCGCAACCCCGAGCTGCAAGACCATCTGAAAGGCGCTGCGCCCGAGCG 2127
QY 1015 GCCAGCTTGAGAGAGATGATGACCGCTGCGAGGGCGTGGGCGGGCCCGCACGCAAGGCC 1074
DB 2128 GCTACCTTAGAGAGAAATGATGACCGCTGCTCAGGAGTGGCGGAGCCCGCGCAAGGCA 2187
```

```
QY 1075 CGCGTGTGCGCGAGGCGATGAGCCAGGCCAACA-----CCAGCGTATGATGACAGAG 1128
DB 2188 CGCGTGTGCGCTGAGGCGCATGAGCCAGGTGACCAACTCCGCTACCATCATATGACAGCGC 2247
QY 1129 AGCAACTTCAAGGCGCGCGCGCGCATGCTCAAGTGTCTTCAACTGCGGCAAGAGGCGCAC 1188
DB 2248 GGCACCTTTCGAAACCAACGCAAGATGCTCAAGTGTCTTCAACTGTGGCAAGAGAGGCGAC 2307
QY 1189 ATGCGCGCAACTGCGCGCGCGCGCGCGCGCAAGAGGCGTGTGCAAGTGTGCGGCAAGAGGCG 1248
DB 2308 ACAGCCCGCAACTGCAAGGCGCGCGCGCGCGCGCGCAAGAGGCGTGTGCAAGTGTGCAAGAGG 2367
QY 1249 CACCAAGATGAAGACTGACCGAGCGCGCGCGCGCAACTTCTGTGGCAAGATCTGGCCCAAGC 1308
DB 2368 CACCAAGATGAAGATTGTACTGAGAGACAGGCTAATTTTAAAGGAAGATCTGGCTTCC 2427
QY 1309 CACAAGGCGCGCGCGCGCGCAACTTCTGTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1368
DB 2428 CACAAGGAGAGCGCGAGGAAATTTCTTCAAGAGCAAGACCAAGCGCGCGCGCGCGCGCGCG 2487
QY 1369 GAGAGCTTCCGCTTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1416
DB 2488 GAGAGCTTCAAGTTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2547
QY 1417 AAGGACCGCGAGACCTTGACCAAGCTGTGAAGAGCGCTGTGGCAAGAGAGAGAGAGAGAG 1476
DB 2548 AAGGACTGTATCTTAACTTCCCTCAGATCACTTGTGGCAGCGAGCGCGCGCTGTACAA 2607
QY 1477 TAA 1479
DB 2608 TAA 2610
```

```
RESULT 14
US-09-872-733A-6
; Sequence 6, Application US/09872733A
; Patent No. 6656706
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as
; TITLE OF INVENTION: MOLECULAR CLONES WITH MUTATED HIV GAG/POL, SIV GAG AND
; TITLE OF INVENTION: SIV ENV GENES
; FILE REFERENCE: 2026-4287US1 HIV GAG/POL, SIV GAG & ENV
; CURRENT APPLICATION NUMBER: US/09/872,733A
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/34985
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/173,036
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 8366
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: Description of Artificial Sequence: DNA sequence
; OTHER INFORMATION: of the construct pCMVgagpolBtkan containing a CMV
; OTHER INFORMATION: promoter, a HIV gag/pol gene and a kanamycin
; OTHER INFORMATION: resistance gene
US-09-872-733A-6
```

```
Query Match 50.3%; Score 744.6; DB 4; Length 8366;
Best Local Similarity 70.5%; Pred. No. 8.5e-105;
Matches 1060; Conservative 0; Mismatches 419; Indels 24; Gaps 4;

QY 1 ATGGGCGCCGCGCGCAGCATCTGCGCGGCGCAAGCTGACGCTGGGAGCGCATCCGC 60
DB 770 ATGGGTGCGAGAGCGTCACTATTAAAGCGGCGGAGAAATTAGATGATGGGAAAAATTCCG 829
QY 61 CTGCGCGCGCGCGCAAGAGTGTCTACATGATGAAGACACTGTGTGGCGAGCGCGAG 120
DB 830 TTAAGGCGAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 889
```



```
QY 121 CTGAGAAAGTTGCGCCCTGTAACCCCGGCTGCTGAGAGACGAGGCGCTGCAAGCAGATC 180
DB 890 CTAGAACGATTCGAGTTAATCTGGCTGTAGAAAATCAGAAAGCTGTAGACAATA 949
QY 181 ATCCGCACTGCAACCCCGCTGCAAGCCGAGGAGCTGAAAGCTGTTCAC 240
DB 950 CTGGGACAGTCAACCATCTCTTCAAGAGATCAGAGAGCTTCGATCACTATAAAC 1009
QY 241 ACCGTGCACTCTGTACTGCTGCAAGAAAGATCGAGTCCGCAACCAAGAGGCC 300
DB 1010 ACAGTAGCAACCTCTTATGTGTGCAAGCGGATCGAGATCAAGGACCAAGGAACT 1069
QY 301 CTGACAAAGATCGAGAGAGAGCAAAAGTTCAGCAAGAGATCCAGCAGCCGAGGCC 360
DB 1070 TTAGACAAGATAGAGAAAGAGCAAAAGTTCAGAAAGAGGCCAGCAGCAGCT 1129
QY 361 GCC-----GACAAAGGCAAGGTGAGCCAGAACTACCCCATCTGCAAGACTGCAAGGC 414
DB 1130 GACACAGGACAGCAATCAGGTCAAGCCAAATTAACCTATATGTCAGAAATCCAGGGG 1189
QY 415 CAGATGTGCAACGAGCCATCAAGCCCGCAACCTGAAAGCTGAGGTGAGTTCAG 474
DB 1190 CAATGTATCATCAGGCTCATATCACTAGAACTTTAAATGATGAGGTAAAGTAGTAA 1249
QY 475 GAGAAAGCTTCAGCCCGAGGTATCCCATGTTCAACCCCTGAGCGAGCGCCACC 534
DB 1250 GAGAAAGCTTCAGCCCGAGGTATACCATGTTTTCAGCATTTATCAGAAAGACCACC 1309
QY 535 CCCCAGACCTGAACAGATGTTGAACACCTGAGGCGGCCACCAAGCGCCATGCAAGT 594
DB 1310 CCACAGACCTGAACAGATGTTGAACACCTGAGGCGGCCACCAAGCGCCATGCAAGT 1369
QY 595 CTGAAGACCATCAACGAGAGGCGCCGAGTGGACCGCTGCAACCCCTGCAAGCC 654
DB 1370 TTAAGAGACCATCAATGAGGAGCTGCAAGATGGATGAGTGCATCCAGTGCATGCA 1429
QY 655 GGGCCCATGCCCCCGGCAAGTGCAGAGCCCGGCGAGCAGCATGCGCGCACACC 714
DB 1430 GGGCCATATGCAACAGGCCAGTGAAGAACCAAGGGAGTGAATAGCAGAACTACT 1489
QY 715 AGCACCTGCAAGAGCAGATCGCTGATGACAGCAACCCCATCCCTGAGGCGAC 774
DB 1490 AGTACCTTCAGAAACAAATGAGATGATGACAAATATCACTATCCAGTAGAGAG 1549
QY 775 ATCTCAAGCGGTGATCATCTGCGCTGGAACAAGATCGTGGATGTATACGCCGTG 834
DB 1550 ATCTCAAGAGGTGATATCTGCGATTTGAACAAGATCGTGGATGTATAGCCCTTACC 1609
QY 835 AGCATCTGCAATCAAGCAGGCGCCCAAGAGCCCTTCGCGACTACGTGACCGCTTC 894
DB 1610 AGCATCTGCAATCAAGCAGGCGCCCAAGAGCCCTTTAGAGACTATGTAGACCGGTTT 1669
QY 895 TTCAAGACCTGCGCGCGAGAGAGACACCAAGAGGTGAAGATCGATGACCGACACC 954
DB 1670 TATAAACTCTAAGAGCTGAGCAAGCTTCACAGAGGTAAATAATGATGACAGAAACC 1729
QY 955 CTGCTGTGAGAAAGCAACCCGACTGCAAGACCATCTGCGCGCTCTCGGCCCGGC 1014
DB 1730 TTGTTGTCTCAAAATGCAACCAAGATTGTAGAACATCTGAAGGCTCTCGGCCAGCG 1789
QY 1015 GCCAGCTGAGAGATGATGACCGCTGCGAGGCGGTGGCGCCCGACCAAGGCC 1074
DB 1790 GCTACACTGAGAAATGATGACAGCATGTGCGGAGTAGAGGACCCGCGCATTAAGCA 1849
QY 1075 CGCGTGTGCGCGAGCGATGAGCCAGCCACA-----CCAGGTGATGATGAGAAG 1128
DB 1850 AGAGTTTGGCGGAGCGATGAGCCAGGTGACGAACCTCGCGCATATATGATGACAGAG 1909
QY 1129 AGCAACTTCAAGGGCGCCCGGCGATGCTCAAGTCTTCAACTGCGCAAGGAGGCCAC 1188
DB 1910 GGCAACTTCCGGAACAGCGGAGATGCTCAAGTCTTCAATGTGCGAAAGAGGCGAC 1969
QY 1189 ATGCGCCGCACTGCGCGCGCCCGCAAGAGGGCTGCTGGAAGTGGCGCAAGGAGGC 1248
```

```
DB 1970 ACCGCCAGAACTGCGGCGCCCGGAGAGAGGCTGTGAAATGTGAAAGAGAGGA 2029
QY 1249 CACCAGATGAAGACTGCACCGAGCGCCAGGCCAACTTCTGCGCAAGATCTGCGCCAGC 1308
DB 2030 CACCAATGAAGATGTACTAGAGACAGGCTAATTTTAAAGGAAGATCTGCGCTTCC 2089
QY 1309 CACAAGGCGCCCGGCAACTTCTGCAAGCGCCCGGAGGCCAAGCCCGCCCGCC 1368
DB 2090 TACAAGGAGGCGGAGGAAATTTCTTCAAGACAGACAGAGCCCAAGCCACAGAA 2149
QY 1369 GAGAGCTTCCCT-----TCGAGAGACCAACCCCGGCGAGAGCAGAG-----AGC 1416
DB 2150 GAGAGCTTCAAGTCTGGGTAGAGACAACTCCCTCAGAGCAGAGCCGATAGAC 2209
QY 1417 AAGACCGCGAGACCTTGAACCACTGAGAGGCTGTTCGCAAGACCCCTGAGCCAG 1476
DB 2210 AAGAACTGTATCTTTAACTTCCCTCAGATCACTTTGGCAAGACCCCTGTCAGAG 2269
QY 1477 TAA 1479
DB 2270 TAA 2272
```

```
RESULT 15
US-09-184-418C-100
; Sequence 100, Application US/09184418C
; Patent No. 6492110
; GENERAL INFORMATION:
; APPLICANT: Hahn, Beatrice
; APPLICANT: Gao, Feng
; APPLICANT: Shaw, George
; TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN
; TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS TYPE 1
; FILE REFERENCE: D6287
; CURRENT APPLICATION NUMBER: US/09/184,418C
; CURRENT FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 112
; SEQ ID NO 100
; LENGTH: 1476
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; OTHER INFORMATION: isolate=94IN476.104; gene=gag
US-09-184-418C-100
```

Query Match 49.9%; Score 738.2; DB 4; Length 1476;  
Best Local Similarity 69.2%; Pred. No. 8e-104;  
Matches 1023; Conservative 0; Mismatches 453; Indels 3; Gaps 1;

```
QY 1 ATGGCGCCCGCGCCAGCATCTGCGCGCGGCAAGCTGAGCGCTGGAGCGCATCCGC 60
DB 1 ATGGGTGCGAGCGCTCAATATTAAGAGGGGGAATAATGATAGATGGGAAAAATTCCG 60
QY 61 CTGCGCCCGCGCGCAAGAGTCTACATGATGAAGCACCTGTGTGGCCAGCCGCGAG 120
DB 61 TTAAGGCCAGGGGAAAGAAACATTATATGATTAACCTTAGTATGCGCAAGAGGGAG 120
QY 121 CTGAGAAAGTTGCGCCCTGAACCCCGGCTGCTGAGAGACCAAGAGGGCTGCAAGCAGATC 180
DB 121 CTGAAAGATTTGCGCTTAACCTGCGCTTTAGAGAGCTCAGACGGATGTAAACAATA 180
QY 181 ATCCGCACTGCAACCCCGCTGCAAGCCGAGCGAGAGAGCTGAAGAGCTGTTCAC 240
DB 181 ATAAACAGCTACATCCAGCTTTAAGACAGAAACAGAGAACTTAGTCAATTAATCAAC 240
QY 241 ACCGTGCGCACCTGTACTGCTGCAAGAGATGAGTCCGCAACCAAGAGGCC 300
DB 241 ACAGTAGCAACTCTTATGTGTGATGATGAGGATGAGTAGAGTGAACCAAGGAAAGCC 300
QY 301 CTGACAAGATGAGAGAGAGAGCAAAAGTTCAGCAAGAGATCCAGAGCGGAGGCC 360
DB 301 TTAGACAAGATGAGAGAGAGCAAAACAAAGTCAAGCAAAATAATACAGAGCGCAAGAGAG 360
```

QY 361 GCCGACAGGGCAAGGTGAGCCAGAACTACCCATCGTGAGAACTGACGGCCAGATG 420  
Db 361 GCTGAC---GAAAGGTGAGTCAAAATATCTATAGTGCAAAATCTCAAGGGCAATG 417  
QY 421 GTGACACGAGGCCATCAGCCCCCGCACCCCTGAACGCTGGGTGAGGTGATGAGAGAAG 480  
Db 418 GTACACGAGCCCCCTATCAGCTAGAACTTGAATGCGTGGTAAAGTAATAGAGAGAAG 477  
QY 481 GCCTTCAGCCCCGAGGTGATCCCATGTTCAACGCCCTGAGCGAGGGCCACCCCCAG 540  
Db 478 GCTTTAGCCCAAGGTATACCCATGTTCAAGCATTTATCAGAAGAGCCACCCCTCT 537  
QY 541 GACCTGAACACGATGTTGAACACCGTGGGGGCCACCAAGCCGCATGCAATGCTGAAG 600  
Db 538 GATTTAACACCATGTTAATAACAGTGGGGGACATCAAGCAGCCATGCAATGTTAAA 597  
QY 601 GACACCATCAACGAGAGGGCCGAGTGGGACCGGTGACCCCGTGACCGCGGCCCC 660  
Db 598 GATACCATCAACGAGAGGGCTCGGGAATGGGATAGATTACATCCAGTACATGACGGCCT 657  
QY 661 ATGCCCCCGCCAGATGCGGAGGCCCGCGGAGCGACATCGCCCGGACCAACGACACC 720  
Db 658 AATCCACGAGGCCAGATGAGAGAACCAAGGGGAAGTATAGCAGAACTACTAGTACC 717  
QY 721 CTGACGAGCAGATCGCTGATGACCAAGCAACCCCATCCCGTGCGGACATCTAC 780  
Db 718 CTTCAAGAACAAATAGCATGATGACAGGTAAACCATCTATCCAGTGGAGACATCTAT 777  
QY 781 AAGCGGTGATCATCTGCGCTGAACAAGATCGTGGGATGTAAGCCCGGTGAGCATC 840  
Db 778 AAAAGATGATATATTCTGGGGTTAATAAATAGTAAGATGTATAGCCCTGTGAGCATT 837  
QY 841 CTGACATCAAGCAGGGCCCCAAGAGGCCCTTCGGAATAAGTGAACCGCTTCTTCAAG 900  
Db 838 TTGGAATTAAGCAAGGGCCAAAGAACCTTTAGAGACTATGTAGACCGGTTCTTTAA 897  
QY 901 ACCCTGCGCCCGGAGCAGAGCAACCCAGAGGTGAAGACTGATGACCGCAACCTGCTG 960  
Db 898 ACTTTAAGAGCTGAACAAGCTACACAGAACTAAAAGTTGATGACAGACACCTGTTG 957  
QY 961 GTGCAAGACGCCCAACCCCGACTGCAAGACCATCTGCGCGCTCTGCGCCCGCGCCAGC 1020  
Db 958 GTCCAAAATGCAAAACCCAGATTGTAGACCATTTTAAGACATTAGACCAAGGGCTTCA 1017  
QY 1021 CTGAGAGAGATGATGACCGCTTGCACAGGCGTGGCGGCCCAAGCCCAAGGCCCGCTG 1080  
Db 1018 TTAGAAGAAATGTGACAGCATGTCAAGAGTGGAGGACTTAGCCACAAAGCAGAGTG 1077  
QY 1081 CTGCGCGAGGCGATGAGCCAGGCCCAACACAGCGTGAATGACAGAAAGCAACTTCAAG 1140  
Db 1078 TTGGCTGAGGCAATGAGCCAAATCAGATGTAACATAATGATGACAGAGGCAATTTTAA 1137  
QY 1141 GGGCCCCGGCGCATCGTCAAGTGTCTTCACTGCGGCAAGAGGGCCACATCGCCCGCAAC 1200  
Db 1138 GGGCTTAAAGAAATTTTAATGCTTCACTGTGCAAGGAAAGGACACATAGCCAGAAAT 1197  
QY 1201 TGGCGGCCCCCGCGAAGAGGGCTGTGGAAGTGCAGAGGGCCACAGATGAAG 1260  
Db 1198 TGCAGGGCCCCCTAGAAAAGAGGCTGTGGAATGTGGGCAAGAGGACCAATGAAA 1257  
QY 1261 GACTGACCCGAGCGCCAGGCCCAACTTCTGGGCAAGATCTGGCCAGCCCAAGGGCGCG 1320  
Db 1258 GACTGTACTGAGAGGCAAGGCTTAATTTTAAAGGAAATTTGGCTTCCCAAGGGGAGG 1317  
QY 1321 CCGGCAACTTCTGCAAGAGCGCCCGGAGCCCAACCGCCCCCGCGGAGAGCTTCCGC 1380  
Db 1318 CCAGGGAATTTCTTCAAAACAGGCCCAAGGCCCAACAGGCCCAAGAGAGCTTCAAG 1377  
QY 1381 TTGAGAGAGACCAACCCCGGCGCAGAGGAGAGAGCAAGGACCGGAGACCTTGACCAAGC 1440  
Db 1378 TTCAAGGAGACCAACCCCGCTCCGAGCGAGAGTGAAGACAGGGAACCTTAACTTCC 1437

QY 1441 CTGAAGAGCTGTTCGGCAACGACCCCTGAGCCAGTAA 1479  
Db 1438 CTCAATCACTCTTTGGCAGCGACCCCTGTCTCAATAA 1476

Search completed: May 28, 2004, 13:26:33  
Job time : 171.017 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 28, 2004, 08:01:50 ; Search time 5783.73 Seconds  
(without alignments)  
7636.275 Million cell updates/sec

Title: US-09-475-704A-3  
Perfect score: 1479  
Sequence: 1 atggcgccgcgcgcagcat.....acgacccctgagccagtaa 1479

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues  
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estln:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hlc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hlc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	ID	Description
C 1	82.6	5.6	746	14	CK215205 FGAS02715
2	81.2	5.5	749	13	BQ744212 WHE4113 A
3	80	5.4	1064	14	CK163121 FGAS01573
C 4	79.8	5.4	1126	14	CK211108 FGAS02294

5	78	5.3	807	14	CB673622	CB673622	OSJNE08D
6	78	5.3	821	14	CB662794	CB662794	OSJNEd07C
7	78	5.3	824	14	CB666192	CB666192	OSJNEd12P
8	78	5.3	851	14	CB645622	CB645622	OSJNEb07H
9	76.6	5.2	807	14	CB673355	CB673355	OSJNE07N
10	76.6	5.2	1138	14	CK163513	CK163513	FGAS01614
11	76.4	5.2	812	14	CB681973	CB681973	OSJNEf08M
12	76.4	5.2	832	14	CB684834	CB684834	OSJNEf14L
13	76.2	5.2	759	10	BF259495	BF259495	HVSMef001
14	75.6	5.1	562	10	BF484304	BF484304	WHE2321_F
15	75.6	5.1	649	14	CB870888	CB870888	HC15J04w
16	75.4	5.1	834	14	CB678700	CB678700	OSJNEf01C
17	75.4	5.1	840	14	CB673814	CB673814	OSJNE081
18	74.8	5.1	766	14	CB629976	CB629976	OSJNEb06K
19	74.8	5.1	797	14	CB658304	CB658304	OSJNEc14E
20	74.8	5.1	810	14	CB618374	CB618374	OSJNEa02D
21	74.8	5.1	823	14	CB627740	CB627740	OSJNEb02O
22	74.8	5.1	826	14	CB641789	CB641789	OSJNEb01J
23	74.8	5.1	838	14	CB659169	CB659169	OSJNEc15O
24	74.8	5.1	865	14	CB628660	CB628660	OSJNEb04F
25	74.6	5.0	708	14	CB878344	CB878344	AZ04_1021
26	74.6	5.0	818	14	CB652454	CB652454	OSJNEc02G
27	74.2	5.0	785	12	BG369139	BG369139	HVSMb1002
28	74.2	5.0	933	28	CC391778	CC391778	PUIHJ707B
29	73.6	5.0	791	14	CB650382	CB650382	OSJNEb14M
30	73.4	5.0	731	13	BQ752847	BQ752847	WHE4119_G
31	72.6	4.9	2598	11	AY103647	AY103647	Zea mays
32	72.4	4.9	648	10	BE517305	BE517305	WHE0616_A
33	72.2	4.9	803	14	CB633391	CB633391	OSJNEb12D
34	72.2	4.9	1308	11	AY104577	AY104577	Zea mays
35	71.8	4.9	764	13	BO804819	BO804819	WHE3559_C
36	71.4	4.8	718	14	CD938268	CD938268	OV.109123
37	71.4	4.8	718	14	CD938289	CD938289	OV.109R04
38	71.2	4.8	782	14	CB661708	CB661708	OSJNEd04P
39	71	4.8	787	14	CB658685	CB658685	OSJNEc14O
40	71	4.8	935	14	CA280077	CA280077	SCVPLB2C0
41	70.8	4.8	2299	11	AY106831	AY106831	Zea mays
42	70.6	4.8	738	13	BU295509	BU295509	603608565
43	70.6	4.8	887	14	CB656981	CB656981	OSJNEc11P
44	70.6	4.8	1146	11	AY104406	AY104406	Zea mays
45	70.4	4.8	784	14	CB650686	CB650686	OSJNEb15E

ALIGNMENTS

RESULT 1  
LOCUS CK215205/c 746 bp mRNA linear EST 09-DEC-2003  
DEFINITION FGAS027158 Triticum aestivum FGAS: Library 6 CAP GATE 1 Triticum  
aestivum cDNA, mRNA sequence.  
ACCESSION CK215205  
VERSION CK215205.1 GI:39621309  
KEYWORDS EST.  
SOURCE Triticum aestivum (bread wheat)  
ORGANISM Triticum aestivum  
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Poideae; Triticeae; Triticum.  
REFERENCE 1 (bases 1 to 746)  
Allard, P., Crosby, W.L., Danyluk, J., Budes, F., Frick, M., Gaudet, D.,  
Genswein, B., Graf, R., Gulick, P., Hrycan, L.D., Laroche, A.,  
Links, M.G., McCarthy, E.L., Monroy, A., Muzak, I., Nilsson, D.,  
Penniket, C., Roach, J.L. and Sarhan, F.  
Functional Genomics of Abiotic Stress In wheat and Canola Crops  
Unpublished (2003)  
Contact: Wm L Crosby  
Bioinformatics  
University of Saskatchewan, Department of Computer Science  
1C101 Engineering Building, 57 Campus Drive, Saskatoon,  
Saskatchewan, S7N 5A9, Canada  
Tel: 306 966 1769  
Fax: 306 966 2033  
TITLE JOURNAL  
COMMENT

Email: fgas\_estes@cs.usask.ca  
This sequence is the direct result of the Base calling software  
phred (default parameters). It is the raw base calls. To aid in the  
identification of the high quality insert the software Lucy  
(default parameters) has been run on this sequence. Lucy identified  
the region [13,662].  
Plate: L6B016 row: K column: 15.

FEATURES  
source  
1. .746  
Location/Qualifiers

/organism="Triticum aestivum"  
/mol\_type="mRNA"  
/db\_xref="taxon:4565"  
/clone\_lib="Triticum aestivum FGAS: library 6 CAP GATE 1"  
/note="Organ: Crown and leaf; Vector: pCMV.SPORT6; Crown  
(50%) and leaf (50%) tissues from wheat cultivar Norstar  
after short exposure times to low temperature in the light  
and in the dark. 12 mRNA populations were combined before  
constructing the library. The first 6 populations: After 7  
days of growth at 20C from wheat cultivar Norstar after  
short exposure times to low temperature in the light and  
in the dark. 12 mRNA populations were combined before  
constructing the library. The first 6 populations: After 7  
days of growth at 20, wheat plants were transferred to 4C  
in the light. 1cm crown sections and green leaf tissue were  
separately harvested after 1, 3, and 6 hours of low  
temperature exposure. The last 6 populations: After 7 days  
of growth at 20C, wheat plants were transferred to 4C in  
the dark. 1cm crown sections and green leaf tissue were  
separately harvested after 1, 3, and 6 hours of low  
temperature exposure. First strand synthesis in this  
library was done in the presence of methylated dCTP  
thereby protecting from internal cleavage with NotI. In  
addition, this library used a primer for second strand  
synthesis that annealed to an artificial sequence (RNA  
oligo) added before first strand synthesis. Therefore when  
sequences from EST generated from this library will be  
masked for vector and adaptor sequences, an additional  
masking step will have to be included to mask this RNA  
oligo that is common to all clones (sequence  
CGACTGGAGCAGGAGCACTGACATGACTGAAGAGTAGAAA)."

ORIGIN

Query Match 5.6%; Score 82.6; DB 14; Length 746;  
Best Local Similarity 48.4%; Pred. No. 0.62;  
Matches 288; Conservative 0; Mismatches 301; Indels 6; Gaps 2;  
124 GAGAGTTGCGCCCTGAGACCCCGGCTGCTGAGACCAAGGCTGCAAGCATCATC 183  
737 GAGATNNCCGAAACACACCATGAACTGATGACGCTCGGCGGAGGCGCATTC 678  
184 CGCCAGCTGACCCCGGCTGAGACCGGCGAGGAGAGCTGAAGAGCTGTTCAACCC 243  
677 CTCAACATGCTGCTCAAGCTCATCGGCGCAAGAGACCATGAGATCGCGTTACACC 618  
244 GTGGCCACCTGTAAGTGTGACGAGAAAGATGATCCGCGACCAAGAGGCGCTG 303  
617 GGGTACTCCCTCTCGCCACCGGCTGCGCATCCCGACGACGCAACATCTTGCCATG 558  
304 GACAAGATGAGAGGAGGAGCAAGTGCCAGAGAGATCCAGAGGCGGAGCGGCC 363  
557 GACATCAACCGGAGAGACTAGAGCTGGGCTGCGCATGAGAGAGCGGCGTGGCG 498  
364 GACAGGGGAGAGTGAAGCAAGTACCCCATGCTGAGAACTGCAAGGCGCATGCTG 423  
497 CACAAGATGACTTCGGGAGGG---CCCGGCGCTGCGGTGAGCGCGCTGCTGAG 441  
424 CACGAGGCGATGAGCGCGGCACTGAAAGCTGGGTGAAGTGAATCGAGAGAGGCC 483  
440 GACGAGGCGCAACGAGCACTTGAAGTCTGTTCTGTGAGCGCGCAAGAGCAACTAC 381  
484 TTCAGCCCGAGGTGATCCCATGTTACCGCGCTGAGCGAGGCGCCACCCCGAGGAC 543  
380 CTCAACTACCAAGAGCGCTCATGAAAGCTGTCAAGCTCGGCGGCTCTCGGCTACGAC 321

QY 544 CTGACACAGATGTTGAACCGTGGCGGCCACAGGCGCCCATGCAAGTCTGAGGAC 603  
DB 320 ---AACACGCTCTGGAACGGCTCCGTGTGCTCCCGCGCGCCCATGCGCAAGTAC 264  
QY 604 ACCATCAACGAGGAGGCGCGGAGTGGAGCGCGGTGACACCCCGTGCAGCGCGCCCATC 663  
DB 263 ATCCGCTACTACCGCGACTTGTCTGACTCAACAAGGCGCTCGCGCGCAAGCGC 204  
QY 664 GCGCCCGCGCATGCGCGAGCGCGCGCGGAGCATCGCGCGCACACACAGCA 718  
DB 203 GTGAGATCTGCCAGCTCCCGCTCGGCGAGCGCATCACCTCTGCGCGCGCA 149

RESULT 2  
LOCUS BQ744212 749 bp mRNA linear EST 17-JUL-2002  
DEFINITION WBR4113\_A01 A01ZS wheat salt-stressed root cDNA library Triticum  
aestivum cDNA clone WBR4113\_A01\_A01, mRNA sequence.  
ACCESSION BQ744212  
VERSION BQ744212.1 GI:21890999  
KEYWORDS EST.  
SOURCE Triticum aestivum (bread wheat)  
ORGANISM Triticum aestivum

REFERENCE  
AUTHORS Anderson, O.D., Akhunov, B., Chao, S., Crossman, C., Deal, K.,  
Dvorak, J., Lazo, G.R., Pham, J., Rausch, C.J., Wilson, C. and Woo, J.  
TITLE The structure and function of the expressed portion of the wheat  
genomes - Salt-stressed root cDNA library  
JOURNAL Unpublished (2002)  
COMMENT Contact: Olin Anderson  
US Department of Agriculture, Agriculture Research Service, Pacific  
West Area, Western Regional Research Center  
800 Buchanan Street, Albany, CA 94710, USA  
Tel: 5105595773  
Fax: 5105595818  
Email: oanderson@pw.usda.gov

Sequences have been trimmed to remove vector sequence and low  
quality sequence with phred score less than 20  
Seq primer: SK primer.  
FEATURES  
source  
1. .749  
Location/Qualifiers

/organism="Triticum aestivum"  
/mol\_type="mRNA"  
/cultivar="Chinese Spring"  
/db\_xref="taxon:4565"  
/clone="WBR4113\_A01\_A01"  
/tissue\_type="Roots"  
/dev\_stage="Full tillering"  
/lab\_host="E. coli SOLR"  
/clone\_lib="wheat salt-stressed root cDNA library"  
/note="Vector: Lambda Uni-ZAP XR, excised phagemid  
pBluescript SK(-); Site 1: BclRI; Site 2: XhoI; Hydroponic  
plants grown to full tillering stage were treated with 150  
mM NaCl for either 12 hours or 7 days. Root tissues of the  
plants subjected to both types of treatment were collected  
separately at University of California, Davis (E. Akhunov  
and K. Deal in J. Dvorak's lab). Total RNA was prepared  
separately from the two samples (12h and 7day treatments),  
and equal amount of RNA was then pooled. PolyA RNA was  
purified from the pooled RNA, a cDNA library was made, and  
the cDNA clones were in vivo excised to give pBluescript  
SK(-) phagemids in J. Dvorak's lab (E. Akhunov, J. Dvorak)  
at the University of California, Davis. Colony plating,  
plasmid DNA preparations and DNA sequencing were performed  
in the OD Anderson lab (all other authors)."

ORIGIN

Query Match 5.5%; Score 81.2; DB 13; Length 749;  
Best Local Similarity 47.5%; Pred. No. 0.93;

	Matches	282;	Conservative	0;	Mismatches	303;	Indels	9;	Gaps	1;
Oy	280	GTCCCGGACACCAGAGAGGCCCTTGACAAGATCGAGGAGAGACAGAACAAGTGCACAG	339							
Db	19	GTTCAGCGCCAAAGCCGGCCGCCCAAGCACCCGCGCTTCGGGCCCAAGTGCGCTTCAACCA	78							
Oy	340	AAGATTCAGCAGGCGGAGGCGCGGCAAGAGGGCAAGGTGAAGCCAATACTAACCATTCTG	399							
Db	79	CGGCCCCAGCTCCCCAGCGCGCGGACGAGAGAGGTGAGAGGCCAAGCGCGGAGGCCGAG	138							
Oy	400	CAGAACCTGCAAGGCGCAGATGTTGACCAAGGCCCATTCAGCCCCCGGCAACCTGAAGCCTTG	459							
Db	139	AAGGAGCGCGCCCGCCCGGAGCGGCACTAACAGAACCAACCGCTTCCACCAACCAAGGCG	198							
Oy	460	GTGAAGGTGATCGAGGAGAGAGGCGCTTTCAGCCCCCGAGGTGATCCCATTGTTACCGCCTG	519							
Db	199	TTCGCGGCGCAACCAAGACAGAACACCGGCCAAGCCCTTCGTGACATTCAGCCCGACTGGA	258							
Oy	520	AGGAGAGGCGCGCCACCCCCCAGAGACTGAACAAGATGTGAACACCGTGGCGGCCACAG	579							
Db	259	ATCTCTGAGCAGATCCCTTCGCGCAACTTCACCAAGCTCTCCTTCGCGGTGCGGACAG	318							
Oy	580	GCCGCCATGCAAGATGCTGAAGGACACCATCAACGAGAGAGGCCGCGCGAGTGGACCGCGTG	639							
Db	319	CCGAGGACCTGCTGCTGTGCGGCGCGCTGACTGCTAGACACCGCGCTTAACGACTGCGTC	378							
Oy	640	CACCCCGTGCACGCGCGGCCCATTCGCCCCCGGCGAGATGCGGAGCCCGCGGAGCGAG	699							
Db	379	AACCCCAAGACGCGCGCGCGCTCGAGCGCTTCAAGAACCGCGCAGTTCTTCAAGATCAC	438							
Oy	700	ATGCGCGGGCACCAAGACACCCCTG-----AGAGCAGATCGCTTGATGACCAAGC	750							
Db	439	ACCAACGACAGACCCCATCATCTCCGCGCTCGCCGAGAGAGACAGGCGCACCGCTTCGCG	498							
Oy	751	MACCCCCCCATCCCCGTGGGCGACATCTACAAGCGGTGATCATCTTGCGCTGAACAAG	810							
Db	499	ACCGACGCCATCTCGCGCGCCCTCATATGACACGCCCCGACGATCTCTCTTGAGACATT	558							
Oy	811	ATCGTGGGANTGTACAGCCCCGTGAGCATCTTGACATCAAGCAGGGGCCCAAG	864							
Db	559	GTCGTGACGCGGCTCGGCACAAAGCTTCTTTCGACAAAGCGCGAGGGCTCCAG	612							
RESULT 3										
CKI63121										
LOCUS										
DEFINITION										
		CKI63121	1064 bp	mRNA	linear	EST 05-DEC-2003				
VERSION		FGAS015739	Triticum aestivum FGAS:	Library 4 Gate 8	Triticum					
KEYWORDS		aestivum cDNA,	mRNA sequence.							
SOURCE		CKI63121								
ORGANISM		CKI63121.1	GI:38993028							
REFERENCE		BST.								
AUTHORS		Triticum aestivum (bread wheat)								
		Triticum aestivum								
		Bakaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;								
		Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;								
		Poideae; Triticeae; Triticum.								
		1 (bases 1 to 1064)								
		Allard,F., Crosby,W.L., Danyluk,J., Eudes,F., Frick,M., Gaudet,D.,								
		Genswein,B., Graf,R., Gulick,P., Hrycan,L.D., Iaroche,A.,								
		Links,M.G., McCarthy,E.L., Monroy,A., Muzak,I., Nilsson,D.,								
		Penniket,C., Roach,J.L. and Sarhan,F.								
		Functional Genomics of Abiotic Stress In Wheat and Canola Crops								
		Unpublished (2003)								
		Contact: Wm L Crosby								
TITLE		Bioinformatics								
JOURNAL		University of Saskatchewan, Department of Computer Science								
COMMENT		1C101 Engineering Building, 57 Campus Drive, Saskatoon,								
		Saskatchewan, S7N 5A9, Canada								
		Tel: 306 966 1769								
		Fax: 306 966 2033								
		Email: fgas_est@cs.usask.ca								
		This sequence is the direct result of the Base calling software								
		Phred (default parameters). It is the raw base calls. To aid in the								

```

identification of the high quality insert the software Lucy
(default parameters) has been run on this sequence. Lucy identified
the region [1,834].
Plate: L4B007 row: C column: 11.
Location/Qualifiers
1..1064
FEATURES
Source
```

```

/organism="Triticum aestivum"
/mol_type="mRNA"
/db_xref="taxon:4565"
/clone_lib="Triticum aestivum FGAS: library 4 Gate 8"
/note="Organ: Crown and leaf; Vector: PCMV.SORT6; Aerial
parts (crown and leaf) of wheat cultivar Norstar during
dehydration stress. 8 mRNA populations were combined
before constructing the library. The first four come from
removing plants from vermiculite (7 day old plants) and
incubating them at 20C on the bench without water for 1,
2, 3 and 4 days. The last four come from plants grown in
soil in a growth chamber after watering is terminated.
Four samplings were taken in a two week period; the first
after wilting was observed and the last, two weeks later,
consisted of live crown and stem tissue (leaf tissue was
yellow and dead). First strand synthesis in this library
was done in the presence of methylated dCTP thereby
protecting from internal cleavage with NotI."

```

**ORIGIN**

Query Match	5.4%;	Score 80;	DB 14;	Length 1064;
Best Local Similarity	48.4%;	Pred. No. 1.4;		
Matches 283; Conservative	0;	Mismatches 296;	Indels 6;	Gaps 21

QY	124	GAGAAGTTGGCCCTGAACCCCGGCTGTGGAGACGAGGAGGGCTGCAGACAGATCATC	183
DB	282	GAGATCACCGCCAAACCAACCATGGAACCTGATGACGACGTGCGCGAGCGAGGGCCAGTTCC	341
QY	184	CGCCAGCTGCACCCCGCCCTGCAGACCGGCGAGCGAGGAGCTGAAGAGCCTGTCAACACC	243
DB	342	CTCAACATGCTGCTCAAGCTCATCGGCGCCAGAAAGATCATGAGATGCGGCTTACACC	401
QY	244	GTGGCCACCCCTGTAATGCGTGCACAGAGAAAGATGAGGTTCGCGACACCAAGAGAGCCCTG	303
DB	402	GGCTACTCCCTCCTCGCCACCGCGCTCGCCATCCCGGACGACGGCACCATCTTGCCATG	461
QY	304	GACCAAGATCGAGAGGAGCAGAAACAATGCCAGCAGAAAGATCCAGCAGGCCGAGCCGCC	363
DB	462	GACATCAACCGCGAAGAACTACAGCTGGGGCTGCCCTGCATCGAAGAGCCGGCTGGCG	521
QY	364	GACAAAGGCAAGGTGAGCCAGAACTACCCCATGTCAGAACTGCAGGGCCAGATGTG	423
DB	522	CACAAAGATGACTTCCGCGAGGG---CCGCGCTGCGGCTGTGACGCGCTGTGAG	578
QY	424	CACCAGGCCATCAGCCCCCGCACCTGAAAGCCTGGGTGAAGGTATCGAGAGAAAGCC	483
DB	579	GACGAGGCCAAACCACGGCACCTTGCACTTGTCTTGTGAGCGCCGACAAGGACCAACTAC	638
QY	484	TTCAAGCCCCGAGGTATCCCCCATGTTCAACCGCCCTGAGCGAGGGCGCCACCCCCAGAC	543
DB	639	CTCAACTACCGAGAGCGCTCTATGAACTGTCTCAAGCTCGGCGGCTCTCGGCTACGAC	698
QY	544	CTGAACACGATGTTGAACAACCGTGGGGCGGCCAACAGGCGCCATGCAATGCTGAAGAC	603
DB	699	---AACACGCTTGGAAACGCTCGCTGTGCTCTCCCGCGAGCGCCCCATGCGCAAGTAC	755
QY	604	ACCATCAACGAGAGGCGCGCTGAGTGGGATCCGCTGCAACCCCGTGCACGCCGCGCCCATC	663
DB	756	ATTCGCTACTACCGCGACTTCTGTCTCTGACCTCAACAAGGCCCTTCGCCGACGACGCGC	815
QY	664	GCCCCCGGCAGATGCGCGAGGCCCGCGCGGCGAGCGCATGCGCGGC	708
DB	816	GTTCGAGATTCGCACTTCCCTCGTGGCGAGCGGCATCACCTCTTGC	860

RESULT 4  
CK211108/c

LOCUS CR211108 1126 bp mRNA linear EST 09-DEC-2003  
 DEFINITION FGAS022942 Triticum aestivum FGAS: Library 6 CAP GATE 1 Triticum  
 aestivum cDNA, mRNA sequence.  
 ACCESSION CR211108  
 VERSION CR211108.1 GI:39617217  
 KEYWORDS EST.  
 SOURCE Triticum aestivum (bread wheat)  
 ORGANISM Triticum aestivum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Poideae; Triticeae; Triticum.  
 1 (bases 1 to 1126)  
 Allard, F., Crosby, W.L., Danyluk, J., Eudes, F., Frick, M., Gaudet, D.,  
 Genswein, B., Graf, R., Gulick, P., Hrycan, L.D., Laroche, A.,  
 Links, M.G., McCarthy, E.L., Montroy, A., Muzak, I., Nilson, D.,  
 Peniket, C., Roach, J.L. and Sarhan, F.  
 Penniket, C., Roach, J.L. and Sarhan, F.  
 Functional Genomics of Abiotic Stress in Wheat and Canola Crops  
 Unpublished (2003)  
 TITLE JOURNAL  
 COMMENT Bioinformatics  
 University of Saskatchewan, Department of Computer Science  
 1C101 Engineering Building, 57 Campus Drive, Saskatoon,  
 Saskatchewan, S7N 5A9, Canada  
 Tel: 306 966 1769  
 Fax: 306 966 2033  
 Email: fgas\_est@cs.usask.ca  
 This sequence is the direct result of the Base calling software  
 phred (default parameters). It is the raw base calls. To aid in the  
 identification of the high quality insert the software Lucy  
 (default parameters) has been run on this sequence. Lucy identified  
 the region [12,750].  
 Plate: L6B001 row: A column: 20.  
 Location/Qualifiers  
 1. 1126  
 /organism="Triticum aestivum"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4565"  
 /clone\_lib="Triticum aestivum FGAS: Library 6 CAP GATE 1"  
 /note="Organ: Crown and leaf; Vector: pCMV.SPORT6; Crown  
 (50%) and leaf (50%) tissues from wheat cultivar Norstar  
 after short exposure times to low temperature in the light  
 and in the dark. 12 mRNA populations were combined before  
 constructing the library. The first 6 populations: After 7  
 days of growth at 20C from wheat cultivar Norstar after  
 short exposure times to low temperature in the light and  
 in the dark. 12 mRNA populations were combined before  
 constructing the library. The first 6 populations: After 7  
 days of growth at 20, wheat plants were transferred to 4C  
 in the light. 1cm crown sections and green leaf tissue were  
 separately harvested after 1, 3, and 6 hours of low  
 temperature exposure. The last 6 populations: After 7 days  
 of growth at 20C, wheat plants were transferred to 4C in  
 the dark. 1cm crown sections and green leaf tissue were  
 separately harvested after 1, 3, and 6 hours of low  
 temperature exposure. First strand synthesis in this  
 library was done in the presence of methylated dCTP  
 thereby protecting from internal cleavage with NotI. In  
 addition, this library used a primer for second strand  
 synthesis that annealed to an artificial sequence (RNA  
 oligo) added before first strand synthesis. Therefore when  
 sequences from EST generated from this library will be  
 masked for vector and adaptor sequences, an additional  
 masking step will have to be included to mask this RNA  
 oligo that is common to all clones (sequence  
 CGACTGAGCAGCAGCACTGACTGACTGAGAGTAGAAA)."

## ORIGIN

Query Match 5.4%; Score 79.8; DB 14; Length 1126;  
 Best Local Similarity 48.2%; Pred. No. 1.5;  
 Matches 287; Conservative 0; Mismatches 302; Indels 6; Gaps 2;

124 GAGAGTTCGCTGACCCCGCTGCTGAGACGAGCGGCTGCAAGCATCATC 183  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 763 GAGATCACCGCCCAACCAACCAATGAACCTGATGACGAGTCGCGGACGAGGCCAGTTTC 704  
 Oy 184 CGCCAGCTGACCCCGCCCTGACAGCCGGACGAGAGCTGAAGACCTGTTCAACACC 243  
 Db 703 CTCAACATGCTGCTCAAGCTCATCGCGCCAGAAAGACCATGAGATCGCGTCTACACC 644  
 Oy 244 GTGGCCACCTCTACTGCGTGCACGAGAATCGAGTCCGCGACCAAGAGGCCCTG 303  
 Db 643 GGCTACTCCCTCTCTCGCCACCGCGCTCGCCATCCCCGACGAGCCACCATCTGGCCATG 584  
 Oy 304 GACCAAGATCGAGAGAGAGAGAAACAATGCGACGAAGAATCCAGAGCCGAGCCGCC 363  
 Db 583 GACATCAACCGGAGAACTACGAGCTGGGCTGCCCTGTCATGAGAAGGCCGGCTGCGG 524  
 Oy 364 GACCAAGGCGAAGGTGAGCCAGAACTACCCCATGCTGCAAGACCTGACAGGCGCAGATG 423  
 Db 523 CACAAGATCGACTTCGCGGAGGG--CCCGGCGCTGCCGTGTCAGCGCGCTGTCGAG 467  
 Oy 424 CACCAAGCCATCAGCCCCCGCAACCTGAAAGCCTGGGTGAAGTGAAGAGAGAGGCC 483  
 Db 466 GACGAGGCCAACCACGCGACCTTGACCTTGCTTCGTGAGCGCCGACAAAGCAACTAC 407  
 Oy 484 TTCAAGCCCGAGGTGATCCCATGTTCAACGCGCTGAGCGAGGCGCCACCCCAAGAC 543  
 Db 406 CTCAACTACCAAGAGCGGCTTATGAAGCTGCTCAAGCTCGCGGCTCTCTGCTACGAC 347  
 Oy 544 CTGAACACGATGTTGAACACCGTGGCGGCGCCACCAAGGCCCATGACATGCTGAAGAC 603  
 Db 346 --AACACGCTCTGGAACGCGCTCGTGTGCTCCCGCGAGCCCGCCATGCGCAAGTAC 290  
 Oy 604 ACCATCAACGAGAGGCGCGGAGTGGAGCCGCGTGCACCCCGTGCACGCGCGCCATC 663  
 Db 289 ATCCGCTCTCTACCGCGACTTCGCTCTGACCTCAACAAGGCCCTCGCGCCGACGCGC 230  
 Oy 664 GCCCCCGCCAGATCGCGAGCGCCCGCGGACGACATCGCGCCACCAACAGCA 718  
 Db 229 GTGAGATCTGCCAGCTCCCGCTGGCGAGCGCATCACCTCTGCGCGCGGCCA 175

## RESULT 5

CB673622

## LOCUS

OSJNB08D20.f OSJNB Oryza sativa (japonica cultivar-group) cDNA

## DEFINITION

clone OSJNB08D20 5', mRNA sequence.

## ACCESSION

CB673622

## VERSION

CB673622.1 GI:29677347

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

Oryza sativa (japonica cultivar-group)  
 Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.  
 1 (bases 1 to 807)  
 Jantasuriyarat, C., Lu, G., Gowda, M., Hatfield, J., Zhou, B., Mazur, E.,  
 Kudrna, D., Dean, R., Soderlund, C., Wang, R. and Wang, G.  
 Large-scale identification of ESTs involved in the interaction  
 between rice and Magnaporthe grisea  
 Unpublished (2003)  
 Contact: Rod Wing  
 Arizona Genomics Institute  
 University of Arizona  
 Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ  
 85721-0088, USA  
 Tel: 520 626 3967  
 Fax: 520 621 9288  
 Email: http://genome.arizona.edu

## PCR Primers

## FORWARD:

gta aac cga cgg cca gtc

## BACKWARD:

gga aac agc tat gac cat g

Plate: 08 row: D column: 20

Seq primer: gta aac cga cgg cca gtcg.

Location/Qualifiers

1. 807

/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="OSJNB08D20"  
/tissue\_type="Leaf"  
/dev\_stage="3 week"  
/lab\_host="DH10B"  
/clone\_lib="OSJNB"  
/note="Vector: pBluescript II KS +; Site\_1: EcoRI; Site\_2:  
XhoI; 24 hrs after inoculation with Rice Blast (70-15)"

ORIGIN

Query Match 5.3%; Score 78; DB 14; Length 807;  
Best Local Similarity 44.4%; Pred. No. 2.3;  
Matches 312; Conservative 0; Mismatches 390; Indels 0; Gaps 0;  
OY 204 GCAGACCGGCGAGGAGCTGAAGAGCTGTTCAACACCGTGGCCACCTGTACTGCGT 263  
DB 90 GCAGATAGAAATGGCGCGGAGACGTTCTTTCACCTCCGAGTCCGTGAACGAGGCTCA 149  
OY 264 GCACGAGAGATCGAGGTCCCGGACACCAAGAGGCGCTGGACAAAGATCGAGAGAGCA 323  
DB 150 CCGGACAAAGCTGTGCGACCAAGGTGTGAGCGCGGTGCTCGACGCGTCCCTCCCAAGA 209  
OY 324 GAACAAGTCCGACGAGAGATTCAGCAGGCGCGGCGCGGCGGACAGGGCAAGGTGAGCCA 383  
DB 210 CCCCAGACAGAGGTGGCGTGGAGACGTGACCAAGACCAACATGTGTGTGTTGCG 269  
OY 384 GAACTACCCCATCTGTGACAGACTGACAGGCGGACATGTGTGACAGGCGCATACGCCCG 443  
DB 270 CGAGATCACCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 329  
OY 444 CACCTGAAAGCGCTGGTGAAGGTGATCGAGAGAGAGGCGCTTCAAGCGCGAGGTGATCC 503  
DB 330 CATCGGCTTGTGTCCGACGACGTGCGGCTTCAAGCGCGGCGGCGGCGGCGGCGGCGG 389  
OY 504 CATGTTACCGCGCTGAGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 563  
DB 390 CATCGAGCAGAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 449  
OY 564 CGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 623  
DB 450 CGAGAGATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 509  
OY 624 CGAGTGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 683  
DB 510 CGAGCTGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 569  
OY 684 GCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 743  
DB 570 CCGCAAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 629  
OY 744 GACCAAGCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 803  
DB 630 GTACCTCAAGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 689  
OY 804 GAACAAGATCGTGGAGTGAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 863  
DB 690 CCAGACGACGAGACCGTCAACCAAGAGAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 749  
OY 864 GAGCGCTTCCGCGGCTACGTTGAGCGGCTTCTTCAAGACCTT 905  
DB 750 CAAGCGGTATCCCGCAAGTAAGTCAAGAGAGACCAT 791

RESULT 6  
CB662794 821 bp mRNA linear EST 09-APR-2003  
LOCUS OSJNB07C09.f OSJNB Oryza sativa (japonica cultivar-group) cDNA  
DEFINITION clone OSJNB07C09 5', mRNA sequence.  
ACCESSION CB662794  
VERSION CB662794.1 GI:29666519

KEYWORDS EST.  
SOURCE Oryza sativa (japonica cultivar-group)  
ORGANISM Oryza sativa (japonica cultivar-group)  
REFERENCE Jantassuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,B.,  
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.  
Large-scale identification of ESTs involved in the interaction  
between rice and Magnaporthe grisea  
Unpublished (2003)  
CONTACT: Rod Wing  
JOURNAL Arizona Genomics Institute  
COMMENT University of Arizona  
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ  
85721-0088, USA  
Tel: 520 626 3967  
Fax: 520 621 9288  
Email: http://genome.arizona.edu  
PCR Primers  
FORWARD: gta aac cga cgg cca gtc  
BACKWARD: gga aac agc tat gac cat g  
Plate: 07 row: C column: 09  
Seq primer: gta aac cga cgg cca gtc.  
Location/Qualifiers  
1..821  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="OSJNB07C09"  
/tissue\_type="Leaf"  
/dev\_stage="3 week"  
/lab\_host="DH10B"  
/clone\_lib="OSJNBd"  
/note="Vector: pBluescript II KS +; Site\_1: EcoRI; Site\_2:  
XhoI; 24 hrs after inoculation with Rice Blast (C9240-1)"

FEATURES

source

ORIGIN

Query Match 5.3%; Score 78; DB 14; Length 821;  
Best Local Similarity 44.4%; Pred. No. 2.3;  
Matches 312; Conservative 0; Mismatches 390; Indels 0; Gaps 0;  
OY 204 GCAGACCGGCGAGGAGCTGAAGAGCTGTTCAACACCGTGGCCACCTGTACTGCGT 263  
DB 88 GCAGATAGAAATGGCGCGGAGACGTTCTTTCACCTCCGAGTCCGTGAACGAGGCTCA 147  
OY 264 GCAGAGAGATCGAGTCCGCGACCAAGAGGCGCTGACAGAGATCGAGAGAGCA 323  
DB 148 CCGGACAAAGCTGTGCGACGAGGTGCGACGCGGTGCTGACGCGTGCCTGCGCAGGA 207  
OY 324 GAACAAGTCCGACGAGATCCAGCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 383  
DB 208 CCGGACAGCAGAGGTGGCGTGCAGACGTGACCAAGACCAAGATGTGTGTTGCG 267  
OY 384 GAACTACCCCATCTGTGACAGACTGACAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 443  
DB 268 CGAGATCACCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 327  
OY 444 CACCTGAAAGCGCTGGTGAAGGTGATCGAGAGAGGCGCTTCAAGCGCGGAGTATCC 503  
DB 328 CATCGGCTTGTGTCCGACGACGTGCGGCTGACGCGCGGCGGCGGCGGCGGCGGCGGCGG 387  
OY 504 CATGTTACCGCGCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 563  
DB 388 CATCGAGCAGAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 447  
OY 564 CGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 623  
DB 448 CGAGAGATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 507  
OY 624 CGAGTGGAGCGGCGTGCACCGCGTGCACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 683



|||||  
DB 508 CGAGCTGATGCCCCCTCAGCCACGCTCTCGCCACCAGCTCGGCGCCGCTCACCAGAGT 567  
OY 684 GCGCCGCGGAGAGCATGCGCCGACACACGACCTCTGACAGAGCAGATGCGCTGGAT 743  
DB 568 CCGCAAGAACGGGCACTGCGCTGCTCAGGCGCCGACGGAAGACCAAGTCAACGTTGA 627  
OY 744 GACCAGCAACCCCCCATCCCGTGGGCGACATCTACAAGCGGTGATCATCTGGGCT 803  
DB 628 GTACTCTCAACGAGCGCCGCGCCATGCTCCCGCTCCACACCGCTCATCTCCAC 687  
OY 804 GAACAAGATCGTGGGATGTACAGCCCCGTGAGCATCTGACATCAAGAGGCGCCCA 863  
DB 688 CCAAGCAGACGAGACCGTCAACCAAGAGATGCGCCGACCTCAAGAGGACGTCAT 747  
OY 864 GGAGCCCTTCCGCGACTACGTGACCGCTTCTCAAGACCT 905  
DB 748 CAAGCCGGTCACTCCCGACAAAGTACTCGACGAGAAAGACAT 789

RESULT 7  
CB666192 824 bp mRNA linear EST 09-APR-2003  
LOCUS OSJNBd12p19.f OSJNB Oryza sativa (japonica cultivar-group) cDNA  
DEFINITION clone OSJNBd12p19 5', mRNA sequence.

CB666192  
CB666192.1 GI:29669917

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Oryza sativa (japonica cultivar-group)  
Oryza sativa (japonica cultivar-group)  
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehretioideae; Oryzae; Oryza.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
1 (bases 1 to 824)  
Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,  
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.  
Large-scale identification of ESTs involved in the interaction  
between rice and Magnaporthe grisea  
Unpublished (2003)  
Contact: Rod Wing  
Arizona Genomics Institute  
University of Arizona  
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ  
85721-0088, USA  
Tel: 520 626 3967  
Fax: 520 621 9288  
Email: http://genome.arizona.edu

PCR Primers  
FORWARD: gta aaa cga cgg cca gtc  
BACKWARD: gga aac agc tat gac cat g  
Plate: 12 row: P column: 19  
Seq primer: gta aaa cga cgg cca gtc.

FEATURES  
source  
1..824  
Location/Qualifiers

/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="OSJNBd12p19"  
/issue\_type="leaf"  
/dev\_stage="3 week"  
/lab\_host="DH10B"  
/clone\_lib="OSJNBd"  
/note="Vector: pBluescript II KS +; Site\_1: EcoRI; Site\_2:  
XhoI; 24 hrs after inoculation with Rice Blast (C9240-1)"

ORIGIN

Query Match 5.3%; Score 78; DB 14; Length 824;  
Best Local Similarity 44.4%; Pred. No. 2.3;  
Matches 312; Conservative 0; Mismatches 390; Indels 0; Gaps 0;

OY 204 GCAGACCGGCGAGGAGAGCTGAAGAGCTGTCAACACCGTGCCACCTGTACTGCGT 263  
|||||

DB 98 GCAGATAGAAATGCGCGGAGAGCTTCTTCACTCCGAGTCCGTGAACAGGCTCA 157  
OY 264 GCACGAGAAGATGAGGTCCCGCACACCAAGAGGCGCTTGACAAAGATCGAGAGGACA 323  
DB 158 CCGGACAAGCTGTGCGACCAAGTGTGGAACGCGGTGCTCGACGCTGCTCGCCAGGA 217  
OY 324 GAACAAGTCCAGAGAGATCCAGCAGGCGCGGCGCCGACAAAGGCAAGTGAAGCA 383  
DB 218 CCGGACAAGAGGTGCGGTGCGAGACGTGACCAAGACCAAGATGTGTATGTTCGG 277  
OY 384 GAACTACCCATGCTGCAAGACCTGCAAGGCGCGAGATGTTGACCAAGGCCATCAGCCCCG 443  
DB 278 CGAGATCACCAAGGCGCACCGTGACTACGAGAAATGCTCCGCGACACCTGCCGCG 337  
OY 444 CACCTGAACGCTGGGTGAAGTGAATGAGAGAGAGGCTTCAAGCCCCGAGTATCCC 503  
DB 338 CATCGGCTTGTGTCCGACGAGCTGGGCTTCGACGCGCCGACCGCTGCAAGGTGCTCA 397  
OY 504 CATGTTCAACGCGCTGAGCGGAGGCGCCACCCCGAGGACCTGAACACGATGTTGAAC 563  
DB 398 CATGAGCAGACAGTCCCGCCGACATCGCGCAGGGGTGCAAGGCCACTTCACCAAGCGCCC 457  
OY 564 CGTGGCGGCGCACAGGCGCGCCATGCAAGTCTGAAGAGACACCATCAACGAGAGCGCG 623  
DB 458 CGAGAGATGCGCGCGCGCGACAGGCGCCACATGTTTCGCTACGCGCACCGAGACCCC 517  
OY 624 CGAGTGGAGCCGCTGCACCCCGTGACGCGCGCCATCGCCCCCGGCGCAAGTGGCGA 683  
DB 518 CGAGCTGATGCGCTTCAGCCACGCTCTCGCCACCAAGCTCGGCGCTCACCAGAGT 577  
OY 684 GCGCCGCGGAGGACATGCGCGCGACACCAAGCCTGCAAGAGCAGATGCGCTGGAT 743  
DB 578 CCGAAGAACGGCACTGCGCTGAGCTCAGGCGCGACGCGCAAGACCAAGTCAACCGTTGA 637  
OY 744 GACCAGCAACCCCCCATTCGCCGTGGGCGACATCTACAAGCGGTGATCATCTGGGCT 803  
DB 638 GTACTCAACGAGCGCGCGCCATGCTCCCGCTGCTCCACACCGTCTCATCTCCAC 697  
OY 804 GAACAAGATCGTGGGATGTACAGCCCCGTGAGATCTTGACATCAAGAGGCGCCCA 863  
DB 698 CCAAGCAGACGAGACCGTCAACCAAGAGATGCGCGCGACCTCAAGAGACGTCAT 757  
OY 864 GGAGCCCTTCCGCGACTACGTGACCGCTTCTCAAGACCT 905  
DB 758 CAAGCCGGTCACTCCCGACAAAGTACTCGACGAGAAAGACAT 799

RESULT 8  
CB645622 851 bp mRNA linear EST 08-APR-2003  
LOCUS OSJNB07H03.f OSJNB Oryza sativa (japonica cultivar-group) cDNA  
DEFINITION clone OSJNB07H03 5', mRNA sequence.

CB645622  
CB645622.1 GI:29640613

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Oryza sativa (japonica cultivar-group)  
Oryza sativa (japonica cultivar-group)  
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehretioideae; Oryzae; Oryza.

REFERENCE  
AUTHORS  
TITLE  
Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,  
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.  
Large-scale identification of ESTs involved in the interaction  
between rice and Magnaporthe grisea  
Unpublished (2003)  
Contact: Rod Wing  
Arizona Genomics Institute  
University of Arizona  
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ  
85721-0088, USA  
Tel: 520 626 3967  
Fax: 520 621 9288

JOURNAL  
COMMENT



Email: <http://genome.arizona.edu>

# PCR Primers

FORWARD: gta aac cga cgg cca gtcg

BACKWARD: gga aac agc tat gac cat g

Plate: 07 row: H column: 03

Seq primer: gta aac cga cgg cca gtcg.

Location/Qualifiers

1..851

/organism="Oryza sativa (japonica cultivar-group)"

/mol\_type="mRNA"

/cultivar="nipponbare"

/db\_xref="taxon:39947"

/clone="OSJNB07H03"

/tissue\_type="Leaf"

/dev\_stage="3 week"

/lab\_host="DH10B"

/clone\_lib="OSJNB"

/note="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2: XhoI; 24 hrs after inoculation with Rice Blast (Che 86061)"

## ORIGIN

Query Match 5.3%; Score 78; DB 14; Length 851;  
Best Local Similarity 44.4%; Pred. No. 2.4;  
Matches 312; Conservative 0; Mismatches 390; Indels 0; Gaps 0;

```

OY 204 GCAGACCGGCGAGAGAGCTGAAGAGCTGTTCACACCGGTGGCCACCTGTACTGCGT
DB 105 GCAGATAGAAATGCGCGCGAGAGAGCTTCTTCACTCCGAGTCCGTGAACGAGGGTCA
OY 264 GCACGAGAGATCGAGGTCCCGGACACCAAGAGAGGCGCTGGACAAGATCGAGGAGGCA
DB 165 CCGGAGAGAGTGTGCGGACCAAGGTGTGGAAGCGGTGCTGAGAGCGTCTGCCCAAGGA
OY 324 GAACAAAGTCCAGCAGAAAGATCCAGAGGCGGAGGCGCGGACCAAGGGCAAGTGAGCCA
DB 225 CCGGAGAGAGTGTGCGGAGAGAGAGTCCAGAGGCGGACCAAGGCGGACCAAGTGAGTGTGCG
OY 384 GAATTAACCCATCTGTGAGAACTGCAAGGCGGAGAGTGTGCAACAGGCGGATCAGCCCGG
DB 285 CGAGATCAACCAAGGCGGAGAGTGTGCAAGAGATGTCGCGGACCACTGCGCGCGG
OY 444 CACCTGAAAGCTGTGAGAGTGTGAGAGAGAGAGGCTTCAGCCCGGAGGTGATCCC
DB 345 CATGGCTTGTGTCCGAGAGAGTGTGAGAGGCGGAGAGGCGGAGAGTGTGTCACAA
OY 504 CATGTTCAACCGCTGAGAGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGG
DB 405 CATGAGCAGCAGTGTGCGGAGAGAGTGTGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGG
OY 564 CGTGGCGGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGG
DB 465 CGAGAGATCGGCGGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGG
OY 624 CGAGTGGAGAGCGGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGG
DB 525 CGAGTGTATGCGGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGG
OY 684 GCGGCGGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGG
DB 585 CCGGAGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGG
OY 744 GACCAAGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGG
DB 645 GTACTCAACGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGG
OY 804 GAACAGAGATGTGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGG
DB 705 CGAGAGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGG
OY 864 GAGGCGGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGG
DB 765 CAGGCGGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGG

```

## RESULT 9

CB673355

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (2003)

Arizona Genomics Institute

University of Arizona

Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ

85721-0088, USA

Tel: 520 626 3967

Fax: 520 621 9288

Email: <http://genome.arizona.edu>

PCR Primers

FORWARD: gta aac cga cgg cca gtcg

BACKWARD: gga aac agc tat gac cat g

Plate: 07 row: N column: 10

Seq primer: gta aac cga cgg cca gtcg.

Location/Qualifiers

1..807

/organism="Oryza sativa (japonica cultivar-group)"

/mol\_type="mRNA"

/cultivar="nipponbare"

/db\_xref="taxon:39947"

/clone="OSJNB07H10"

/tissue\_type="Leaf"

/dev\_stage="3 week"

/lab\_host="DH10B"

/clone\_lib="OSJNB"

/note="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2: XhoI; 24 hrs after inoculation with Rice Blast (70-15)"

ORIGIN

Query Match 5.2%; Score 76.6; DB 14; Length 807;

Best Local Similarity 44.3%; Pred. No. 3.5;

Matches 310; Conservative 0; Mismatches 389; Indels 0; Gaps 0;

```

OY 204 GCAGACCGGCGAGAGAGCTGAAGAGCTGTTCACACCGGTGGCCACCTGTACTGCGT
DB 109 GCAGATAGAAATGCGCGGAGAGAGCTTCTTCACTCCGAGTCCGTGAACGAGGGTCA
OY 264 GCACGAGAGATCGAGGTCCCGGACACCAAGAGAGGCGCTGGACAAGATCGAGGAGGCA
DB 169 CCGGAGAGAGTGTGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGG
OY 324 GAACAGTGTCCAGCAGAAAGATCCAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGG
DB 229 CCGGAGAGAGTGTGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGG
OY 384 GAATTAACCCATCTGTGAGAACTGCAAGGCGGAGAGTGTGCAACAGGCGGATCAGCCCGG
DB 289 CGAGATCAACCAAGGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGG
OY 444 CACCTGAAAGCTGTGAGAGTGTGAGAGAGAGAGGCTTCAGCCCGGAGGTGATCCC
DB 349 CATGGCTTGTGTCCGAGAGAGTGTGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGGAGAGGCGG

```

QY 504 CATGTTACCCGCTGAGCGAGGCGCCACCCCGACCTTGAACAGATTTGAACAC 563  
DB 409 CATGAGCAGCAGTCCGCGGACATGCGGAGGGGTGCACGGCACTTCAACAGCGCCC 468  
QY 564 CTTGGGCGGCGCAGCAGCGCGGATGCTGAAGGACCATCAAGAGAGCGCGC 623  
DB 469 CGAGAGATCGCGCGCGGACGAGGCGCACATGTTGGCTACGCGCACGAGAGACCCC 528  
QY 624 CGAGTGGAGCGCGTGCACCCCGTGCAGCGCGCCCATCGCCCGCGGAGATCGCGA 683  
DB 529 CGAGCTGATGCGCTTCAAGCAGCTCTGCGACCAAGCTCGCGCGCTTCAAGAGT 588  
QY 684 GCGCGCGCGCAGCAGCATCGCGCGCACCGACCGCTGAGAGAGATCGCTGGAT 743  
DB 589 CGGCAAGAACGCGACCTGCGCTGCTGAGCGCGCACGAGAGAGAGAGAGAGAGTGA 648  
QY 744 GACGAGCAGCAGCT 803  
DB 649 GTACCTCAAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTCAAC 708  
QY 804 GAAAGATGCTGGGATGTACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 863  
DB 709 CGAGCAGCAGCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCAT 768  
QY 864 GAGCGCTTCCGCGCTACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 902  
DB 769 CAGCGCGCTCATCCCGCAGCAGTACTCGAGAGAGAG 807

RESULT 10  
CK163513 1138 bp mRNA linear EST 05-DEC-2003  
LOCUS FGAS016142 Triticum aestivum FGAS: Library 4 Gate 8 Triticum  
DEFINITION aestivum cDNA, mRNA sequence.  
ACCESSION CK163513  
VERSION CK163513.1 GI:38993817  
KEYWORDS EST.  
SOURCE Triticum aestivum (bread wheat)  
ORGANISM Triticum aestivum

REFERENCE  
AUTHORS Allard, P., Crosby, W.L., Danyluk, J., Budes, F., Frick, M., Gaudet, D.,  
Genevieve, B., Graf, R., Gulick, P., Hrycan, L.D., Larocque, A.,  
Links, M.G., McCarthy, E.L., Montoy, A., Muzak, I., Nilsson, D.,  
Peniket, C., Roach, J.L. and Sarhan, F.  
Functional Genomics of Abiotic Stress in Wheat and Canola Crops  
Unpublished (2003)  
CONTACT: Wm L Crosby  
COMMENT Bioinformatics  
University of Saskatchewan, Department of Computer Science  
1C101 Engineering Building, 57 Campus Drive, Saskatoon,  
Saskatchewan, S7N 5A9, Canada  
Tel: 306 966 1769  
Fax: 306 966 2033  
Email: fgas\_estes@usask.ca  
This sequence is the direct result of the Base calling software  
phred (default parameters). It is the raw base calls. To aid in the  
identification of the high quality insert the software Lucy  
(default parameters) has been run on this sequence. Lucy identified  
the region [8,827].  
Plate: L4B008 row: F column: 04.  
Location/Qualifiers

FEATURES  
SOURCE 1. 1138  
/organism="Triticum aestivum"  
/mol\_type="mRNA"  
/db\_xref="taxon:4565"  
/clone\_lib="Triticum aestivum FGAS: Library 4 Gate 8"  
/note="Organ: Crown and leaf; Vector: pCMV.SPORT6; Aerial  
parts (crown and leaf) of wheat cultivar Norstar during  
dehydration stress. 8 mRNA populations were combined  
before constructing the library. The first four come from

removing plants from vermiculite (7 day old plants) and  
incubating them at 20C on the bench without water for 1,  
2, 3 and 4 days. The last four come from plants grown in  
soil in a growth chamber after watering is terminated.  
Four samplings were taken in a two week period; the first  
after wilting was observed and the last, two weeks later,  
consisted of live crown and stem tissue (leaf tissue was  
yellow and dead). First strand synthesis in this library  
was done in the presence of methylated dCTP thereby  
protecting from internal cleavage with NotI."

## ORIGIN

Query Match 5.2%; Score 76.6; DB 14; Length 1138;  
Best Local Similarity 49.0%; Pred. No. 3.7;  
Matches 291; Conservative 0; Mismatches 295; Indels 8; Gaps 3;

QY 124 GAGAGTTGCGCTGAAACCCCGGCTGCTGAGACCAAGGCGCTGCAAGCATATC 183  
DB 232 GAGATCACCAGCAACCAACCATGGAACCTGATGACAGCTCGGCGGAGAGGCGCATTC 291  
QY 184 CGCAGCTGACACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 243  
DB 292 CTCAATGCTGCTCAAGCTCATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 351  
QY 244 GTGCGCACCTGTACTGCGTGCAGAGAAAGATCGAGTCCGCGACACCAAGAGCGCTG 303  
DB 352 GGCTACTCCCTGCTCGCCACCGCGCTGCGCATCCCGACGAGCGGACCATCTTGGCATG 411  
QY 304 GACAGATCGAG 363  
DB 412 GACATCAACCGCGAGAACTACGAGCTGGGCTGCGTGCATGAGAGAGCGCGCTGGCG 471  
QY 364 GACAGGCGAG 423  
DB 472 CACAGATGACTTCCGCGAG--GGCGCGCGCTGCGCTGCGAGCGCGCTGCTGAGAG 528  
QY 424 CACAGGCGAG 483  
DB 529 GACAGGCGAG 588  
QY 484 TTGAGCGCGGAG 543  
DB 589 CTCACTACGAG 648  
QY 544 CTGAGCAGAGATGTTGAACACCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 603  
DB 649 ---AACACCTCTGGAACGCGCTCGTGTGCTTCCCGCGAGCGCGCGCGCGCGCGG 705  
QY 604 ACCATCAAG 661  
DB 706 ATCCGCTACTACCGCGAGCTTCTGCTGAGCTCAAGAGCGCGCTGCGCGCGCGCG 765  
QY 662 TCGCGCGCGCGAG 715  
DB 766 TCGAGATCTGCGAGCTCCCGGTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 819

RESULT 11  
CB681973 812 bp mRNA linear EST 09-APR-2003  
LOCUS OSJNEf08M09.f OSJNEf Oryza sativa (japonica cultivar-group) cDNA  
DEFINITION clone OSJNEf08M09 5', mRNA sequence.  
ACCESSION CB681973  
VERSION CB681973  
KEYWORDS EST.  
SOURCE Oryza sativa (japonica cultivar-group)  
ORGANISM Oryza sativa (japonica cultivar-group)  
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE  
AUTHORS Jantasuriyarat, C., Lu, G., Gowda, M., Hatfield, J., Zhou, B., Mazur, B.,  
Kudrna, D., Dean, R., Soderlund, C., Wang, R. and Wang, G.

TITLE Large-scale identification of ESTs involved in the interaction between rice and Magnaporthe grisea  
JOURNAL Unpublished (2003)  
COMMENT Contact: Rod Wing  
Arizona Genomics Institute  
University of Arizona  
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ 85721-0088, USA  
Tel: 520 626 3967  
Fax: 520 621 9288  
Email: http://genome.arizona.edu

PCR Primers  
FORWARD: gta aaa cga cgg cca gtc  
BACKWARD: gga aac agc tat gac cat g  
Plate: 08 row: M column: 09  
Seq primer: gta aaa cga cgg cca gtc.

FEATURES  
source 1..812  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="OSJNBf08M09"  
/issue\_type="Leaf"  
/dev\_stage="3 week"  
/lab\_host="DH10B"  
/clone\_lib="OSJNBf"  
/note="Vector: pBluescript II KS +; Site\_1: EcoRI; Site\_2: XhoI; Uninfected Control"

ORIGIN

Query Match 5.2%; Score 76.4; DB 14; Length 812;  
Best Local Similarity 44.3%; Pred. No. 3.7;  
Matches 311; Conservative 0; Mismatches 391; Indels 0; Gaps 0;

204 GCAGACCGGACGAGAGAGCTGAAGCCTGTTCAACACCGTGCCACCTGTACTGCGT 263  
105 GCAGATAGAAATGGCGGCGAGAGCTTCTTCACTCCGAGTCCGTGAACGAGGTCA 164  
264 GCAGGAGAGATCGAGTCCGCGACACCAAGAGGCCCTGACAGATCGAGAGAGCA 323  
165 CCGGACAAAGCTGTGCGACCAAGGTGTGAGCGCGTGTGACGCGTGTGCTGCCAGGA 224  
324 GAACAAGTGCAGAGATCCAGCAGGCGCGAGCGCCGACAGGCGCAAGGTGAGCCA 383  
225 CCGGACAGAGAGTGGCGTGCAGACGTGACCAAGACCAACATGTGTGTTCGG 284  
384 GAACCTACCCATCGTGCAGAACTGACAGGCGCATGTGTGACCAAGCCATCAGCCCCG 443  
285 CGAGATCACCAACCAAGGCGCACCGTCACTAAGAAATGTGCGCGACACCTGCGCGG 344  
444 CACCCTGAACGCTGGGTGAAGTGTGAGGAGAAAGCCTTTCAGCCCCGAGGTATCCC 503  
345 CATCGGCTTCTGTGTCGACGACGTGCGCTCGACGCGCGACCGCTGCAAGTGTCTCAA 404  
504 CATGTTCAACGCGCTGAGCGAGGCGCGCACCCCGCAGACCTGAACAGATGTGAACAC 563  
405 CATCGAGCAGAGTCCCGGACATTCGCGCAGGGGGTGCACGGCCACTTCAACCAAGCGCC 464  
564 CGTGGCGCGCACCGCGCATGACATGCTGAAGACACCATCAACGAGGAGCGCGC 623  
465 CGAGGAGATCGCGCGCGCGGACGAGGCGCACATGTTCTGCTACGCCACGACGAGACCC 524  
624 CGAGTGGACCGCGTGCACCGCGTGCAGCGCGCGCCCATCGCCCCCGGCGAGATCGCGA 683  
525 CGAGTGAATGCCCTCAAGCATTCCTGCGCACCAAGCTCGCGCGCGCTCAACGAGGT 584  
684 GCGCGCGCGGACGACATCGCGCGGACACCAAGACCTTGCAGAGAGAGATCGCTGAT 743  
585 CCGCAAGAGGACACTGCGCGCTGGCTCAGGCGCGGACGAGCAAGACCAAGTCAACGTTGA 644  
744 GACCAAGACCCCTCCATCCCGTGGGCGACATCTACAAGCGGTGATCATCTGGGCT 803

Db 645 GTACCTCAACGAGCGCGCGCCATGTCCTCCGCGTCCACACCGTCTCATCTCCAC 704  
Qy 804 GAACAGATCGTGGCATGTACAGCCCCGTGAGCATCTTGACATCAAGCAGGCCCAA 863  
Db 705 CCAGACGACGAGACCGTCAACCAAGACAGATCGCGCGCGACCTCAAGAGACGTCAT 764  
Qy 864 GAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACCT 905  
Db 765 CAAGCCGTCTATCCGACAGTACTCTGACGAGAGACCAT 806

RESULT 12  
CB684834 832 bp mRNA linear EST 09-APR-2003  
LOCUS  
DEFINITION OSJNBf14L24.f OSJNBf Oryza sativa (japonica cultivar-group) cDNA  
clone OSJNBf14L24 5', mRNA sequence.  
ACCESSION  
CB684834  
KEYWORDS  
SOURCE  
ORGANISM  
EST.  
Oryza sativa (japonica cultivar-group)  
Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.  
1 (bases 1 to 832)  
Jantassuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,  
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.  
Large-scale identification of ESTs involved in the interaction  
between rice and Magnaporthe grisea  
Unpublished (2003)  
Contact: Rod Wing  
Arizona Genomics Institute  
University of Arizona  
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ 85721-0088, USA  
Tel: 520 626 3967  
Fax: 520 621 9288  
Email: http://genome.arizona.edu

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Query Match 5.2%; Score 76.4; DB 14; Length 832;  
Best Local Similarity 44.3%; Pred. No. 3.7;  
Matches 311; Conservative 0; Mismatches 391; Indels 0; Gaps 0;

204 GCAGACCGGACGAGAGCTGAAGCCTGTTCAACACCGTGCCACCTGTACTGCGT 263  
97 GCAGATAGAAATGGCGGCGAGAGCTTCTTCACTCCGAGTCCGTGAACGAGGTCA 156  
264 GCAGGAGAGATCGAGTCCGCGACACCAAGAGGCCCTGACAGATCGAGAGAGCA 323  
157 CCGGACAAAGCTGTGCGACCAAGGTGTGAGCGCGGTGTGACGCGTGTGCTGCCAGGA 216  
324 GAACAAGTGCAGAGATCCAGCAGGCGCGAGCGCCGACAGGCGCAAGGTGAGCCA 383  
217 CCGGACAGAGAGTGGCGTGCAGAGACGTGACCAAGACCAACATGTGTGTTCGG 276

ORIGIN



Db 487 ATCCGCTACTACCGGACCTTGTGCTGACCTCAACAGAGCCCTCGCCGACAGCCG 546  
 Oy 664 GCGCCCGGACAGATGCGCGAGCCCGCGGACGACATCGCCGCGACACAGCA 718  
 Db 547 GTMGAGATGTGCCAGCTCCCGCTCGGCGAGCGCATCACCTCTGCGCGCGCGCA 601

## RESULT 14

BP484304 562 bp mRNA linear EST 06-DEC-2000  
 LOCUS BP484304

DEFINITION WHE2321\_P12\_K23ZS wheat pre-anthesis spike cDNA library Triticum  
 aestivum cDNA clone WHE2321\_P12\_K23, mRNA sequence.

ACCESSION BP484304  
 VERSION BP484304.1 GI:11567605

KEYWORDS EST.  
 SOURCE Triticum aestivum (bread wheat)  
 ORGANISM Triticum aestivum

REFERENCE Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Poideae; Triticaceae; Triticum.  
 1 (bases 1 to 562)

AUTHORS Anderson, O.D., Chao, S., Choi, D.W., Close, T.J., Fenton, R.D.,  
 Han, P.S., Hsia, C.C., Kang, Y., Lazo, G.R., Miller, R., Rausch, C.J.,  
 Seaton, C.L. and Tong, J.C.

TITLE The structure and function of the expressed portion of the wheat  
 genomes - Pre-anthesis spike cDNA library

JOURNAL Unpublished (2000)  
 COMMENT Contact: Olin Anderson  
 US Department of Agriculture, Agriculture Research Service, Pacific  
 West Area, Western Regional Research Center  
 800 Buchanan Street, Albany, CA 94710, USA  
 Tel: 5105595773  
 Fax: 5105595818

Email: oanderson@pw.usda.gov  
 Sequence has been trimmed to remove vector sequence and low  
 quality sequence with phred score less than 20  
 Seq primer: StrataGene SK primer.

FEATURES  
 source Location/Qualifiers

1..562  
 /organism="Triticum aestivum"  
 /mol\_type="mRNA"  
 /cultivar="Chinese Spring"  
 /db\_xref="taxon:4565"  
 /clone="WHE2321\_P12\_K23"  
 /tissue\_type="Spike before anthesis"  
 /dev\_stage="Adult plant"  
 /lab\_host="E. coli SOLR"  
 /clone\_lib="wheat pre-anthesis spike cDNA library"  
 /note="Vector: Lambda Uni-ZAP XR, excised phagemid;  
 Site 1: EcoRI; Site 2: XhoI; Plants were grown in the  
 greenhouse. Whole spike with awns trimmed, white, green  
 and yellow anther were collected and total RNA, and  
 poly(A) RNA were prepared, a cDNA library was made, and  
 the cDNA clones were in vivo excised to give Bluescript  
 phagemids in the TJ Close lab (Choi, Close, Fenton) at  
 the University of California, Riverside. Plasmid DNA  
 preparations and DNA sequencing were performed in the OD  
 Anderson lab (all other authors)."

## ORIGIN

Query Match 5.1%; Score 75.6; DB 10; Length 562;  
 Best Local Similarity 50.4%; Pred. No. 4.3;  
 Matches 270; Conservative 0; Mismatches 254; Indels 12; Gaps 3;

Oy 182 TCCGCGACGTGACCCCGCTGACAGCCGCGAGCGTGAAGAGCTGTTCACCA 241  
 Db 26 TCCTCAACATGCTGCTCAAGCTCATCGCGCAAGAGCATGAGATCGCGCTTACA 85  
 Oy 242 CCGTGGCCACCTGTACTGCGTGACAGAGATGAGGTCCGACACCAAGAGGCC 301  
 Db 86 CCGGCTACTCCCTGCTGCGACCGCGCTCGCATCCCGACGACGCGACCACTTGGCCA 145

Oy 302 TGGACAAAGATGAGAGAGAGCAAGACAGTCCAGAGATCCAGAGGCCGAGGCC 361  
 Db 146 TGGACATCAACCGCGAGAACTACTATCTGGGGCTGCCGTGCATGAGAAAGCCGCGTGG 205  
 Oy 362 CCGACAAAGGCGAAGTGAAGCCAGAACTACCCCATCTGTGACAACTTGAGGGCCAGATGG 421  
 Db 206 CGACAAAGATGACTTCCCGGAGGG--CCCGGCGCTGCCGCTTGAAGCGCTGCTGG 262  
 Oy 422 TGCACAGGCGCATGAGCCCGCCGACCCCTGAACGCTGGGTGAAGTATGAGAGAAAG 481  
 Db 263 AGAGAGAGGCGCAACCAAGGAGCTTGCAGCTTGTCTTGTGAGCGCCGACAAAGCAACT 322  
 Oy 482 CTTGAGCCCGGAGGTGATCCCATGTTTACCGCCCTGAGCGAGGGCGCCACCCCGCAGG 541  
 Db 323 ACCTCAACTACACGAGCGGCTCATGAAGCTGTCAAGGTGCGCGGCTCTCGGCTACG 382  
 Oy 542 ACCTGAACAGATGTTGAACACCGTGGCGGCAACGAGCCGCGCATGAGATCTGAAG 601  
 Db 383 AC--AACACCTCTGGAACGCGCTCCGTGCTGCTCCCGCGAGCCCGCATGCGCAAGT 439  
 Oy 602 ACACATCAACGAGAGGCGCGCGAGTGAGCGCGTGACCCCGTGACGCGCGCCCA 661  
 Db 440 ACATCCGCTACTA-----CCGCACTTGTCTTGTGAGCTCAACAAAGGCCCTGCGCGCG 493  
 Oy 662 TCGCCCGCGCGAGATGCGCGAGCCCGCGGACGACATCGCGCGCACCAAGC 717  
 Db 494 ACCAGCGCTGAGATCTGCCAGCTCCCGTGGCGAGCGCATCAACCTTGGCGC 549

## RESULT 15

CB870888

LOCUS CB870888 649 bp mRNA linear EST 03-JUL-2003  
 DEFINITION HCl5J04w CH Hordeum vulgare cDNA clone HCl5J04 5-PRIME, mRNA  
 sequence.

ACCESSION CB870888  
 VERSION CB870888.1 GI:30072868

KEYWORDS EST.  
 SOURCE Hordeum vulgare

ORGANISM Hordeum vulgare

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Poideae; Triticaceae; Hordeum.  
 1 (bases 1 to 649)

AUTHORS Varshney, R.K., Zhang, H., Burton, R., Stein, N., Langridge, P. and  
 Graner, A.

TITLE Barley ESTs from coleoptile tissue  
 JOURNAL Unpublished (2003)  
 COMMENT Contact: Stein Nils

Molecular Markers Group, Department Genbank  
 Institute of Plant Genetics and Crop Plant Research (IPK)  
 Corrensstr. 3, 06466, Gatersleben, Germany  
 Tel: 039482-5522  
 Fax: 039482-5595

Email: stein@ipk-gatersleben.de  
 Insert length: 649 Std Error: 0.00  
 Plate: 15 row: J column: 4  
 Seq primer: T7.

FEATURES  
 source Location/Qualifiers

1..649  
 /organism="Hordeum vulgare"  
 /mol\_type="mRNA"  
 /cultivar="Sloop"  
 /db\_xref="GABI:549289"  
 /db\_xref="taxon:4513"  
 /clone="HCl5J04"  
 /tissue\_type="coleoptile"  
 /dev\_stage="coleoptile, 3 days old"  
 /lab\_host="DH10B"  
 /clone\_lib="CH"  
 /note="Vector: pSPORT, Site 1: SalI (5-end of cDNA);  
 Site 2: NotI (3-end of cDNA); Due to the cloning system  
 used blue/white selection for recombinants is not 100%  
 reliable. Average insert size is 1.3 kb."



## ORIGIN

Query Match 5.1%; Score 75.6; DB 14; Length 649;  
Best Local Similarity 47.7%; Pred. No. 4.4;  
Matches 284; Conservative 0; Mismatches 305; Indels 6; Gaps 2;

```
OY 124 GAGAAAGTTGCGCCCTGAACCCCGGCTGTGAGACGAGGGCTGCAAGCAGATCATC 183
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 45 GAGATCACCGCAACCAACCATGAACTGTATGACGTGCGGAGAGAGGGCCAGTTG 104
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 184 CGCCAGCTGCACCCCGCCCTGCAAGACCGGAGGAGAGAGCTGAAAGAGCTGTTCAAC 243
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 105 CTCAACATGCTGTCAAGCTCATGCGCGCCAAAGAACCATGAGATCGGCGTTACACG 164
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 244 GTGGCCACCTGTACTGCGTGACAGAAAGATCGAGGTCCGCGACCAAGGAGGCTG 303
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 165 GGTACTCTCCCTGTGCGCACCGCGCTGCGCATCCCGACGACGGCACATCTTGCCATG 224
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 304 GACAAGATCGAGAGAGAGCAAGACAGTGCCAGCAGAAAGATCCAGCAGGCGCGAGCGCC 363
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 225 GACATCAACCGCGAAGACTACAGGTGGGGCTGCGTGCATCGAAGAGCGCGCTGCGG 284
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 364 GACAAGGGCAAGGTGAGCCAGAACTAACCCCATCTGTCAGAACTGCAAGGCGCCAGATG 423
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 285 CACAAGATGACTTCCGCGAGGGCGCGCGCTC---CCCGTCTGACGCGCTCTCTGAG 341
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 424 CACAGGCGCATGACCCCGCACCTCTGAACGCTGCGTGAAGGTGATCGAGAGAGGCGC 483
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 342 GACGAGGCCAACCAAGGCACTTGCATCTTCTGTGACGCCGACCAAGGACAACTAC 401
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 484 TTCAGCCCCGAGGTGATCCCCCATGTTCAACGCTGAGCGAGGCGCCACCCCGCAGGAC 543
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 402 CTCACACTACACGAGCGCTCATGAGCTCGTCAAGGTGCGCGGCTCTCTCGGCTACGAC 461
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 544 CTGAACACGATGTTGAACACCGTGGCGCGCCACGAGCGCCCATGCAATGCTGAAGGAC 603
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 462 ---AACACCTCTGGAACGCTCCGTGTGCTCCCGCGAGCGCCCATGCGCAAGTAC 518
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 604 ACCATCAACGAGAGGCGCGCGAGTGGGACCGCGTGCACCCCGTGCACGCGGCGCATC 663
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 519 ATCGGCTACTACCGCGACTTGTCTGCACTCAACAAAGCCCTGCGCGAGACCAAGCGC 578
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 664 GCCCGCGGCAAGTGCAGAGCGCGCGCGGCGAGCATGCGCGGACCAACCAAGCA 718
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 579 GTGAGATCTGCGCAGCTCCCGCGGCGAGCGCATACCTCTGCGCGCGCGCA 633
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: May 28, 2004, 11:33:43  
Job time : 5789.73 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

CM nucleic - nucleic search, using sw model

Run on: May 28, 2004, 08:01:50 ; Search time 234.634 Seconds  
(without alignments)  
7636.275 Million cell updates/sec

Title: US-09-475-704A-2

Perfect score: 60  
Sequence: 1 gacatccgcagggccccc...tggaacgcttctcaagacc 60

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vit:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	29.8	49.7	646	29	CC821047 FECP4B6 U
2	29.8	49.7	686	14	CA180823 SCACST316
3	29.6	49.3	505	28	BH786628 fzmD014f0
4	29	48.3	1151	13	BQ619287 RMOSE05H1

5	29	48.3	1151	13	BQ619288	BQ619288 RMOSE06A0
6	29	48.3	1151	13	BQ619295	BQ619295 RMOSE06A0
7	29	48.3	1151	13	BQ619376	BQ619376 RMOSE07A0
8	29	48.3	1151	13	BQ619455	BQ619455 RMOSE07H0
9	29	48.3	1151	13	BQ619551	BQ619551 C11lastPC
10	28.6	47.7	709	13	CA072199	CA072199 SCCCAM100
11	28.6	47.7	718	12	BG647389	BG647389 ESTS09008
12	28.4	47.3	525	12	BG549181	BG549181 947073D08
13	28.4	47.3	544	14	CA210513	CA210513 SCEPSB112
14	28.4	47.3	577	29	CG250176	CG250176 OGVCM47TV
15	28.4	47.3	580	12	BG550427	BG550427 947074H04
16	28.4	47.3	744	29	CG226900	CG226900 OGMIM66TH
17	28.4	47.3	592	12	BG550427	CG327425 OGYDB76TV
18	28.4	47.3	857	29	CG327425	CG327425 OGYDB76TV
19	28.2	47.0	408	14	CB639810	CB639810 OSJNBa12B
20	28.2	47.0	576	29	CG225792	CG225792 OGYDB76TV
21	28.2	47.0	576	29	CG225799	CG225799 OGYDB76TV
22	28.2	47.0	585	10	AM147097	AM147097 707011P07
23	28.2	47.0	597	29	CG225799	CG225799 OGYDB76TV
24	28.2	47.0	675	29	CG249691	CG249691 OGYDB76TV
25	28.2	47.0	819	29	CG293655	CG293655 OGYDB76TV
26	28.2	47.0	819	29	CG349088	CG349088 OGYBL67TV
27	28.2	47.0	823	29	CG288932	CG288932 OGYCL45TV
28	28.2	47.0	829	29	CG682396	CG682396 OGYLB77TV
29	28.2	47.0	829	29	CG313210	CG313210 OGYDB76TV
30	28.2	47.0	858	29	CG364851	CG364851 OGYDB76TV
31	28.2	47.0	869	29	CG707427	CG707427 OGYDB76TV
32	28.2	47.0	923	29	CG731054	CG731054 OGYDB76TV
33	28	46.7	500	14	CA645640	CA645640 wreln.pk0
34	28	46.7	524	12	BG313886	BG313886 WHE2066_B
35	28	46.7	615	14	CD905733	CD905733 P468.102L
36	28	46.7	615	28	CC415368	CC415368 P468.102L
37	28	46.7	629	14	CA731056	CA731056 wreln.pk0
38	28	46.7	647	10	BF473271	BF473271 WHE0926_E
39	28	46.7	656	29	CG709393	CG709393 1119012H1
40	28	46.7	656	29	CG310514	CG310514 OGYDB76TV
41	28	46.7	710	14	CD932953	CD932953 GR45.119H
42	28	46.7	763	28	BZ813699	BZ813699 PUGFA26TB
43	28	46.7	777	29	CG240255	CG240255 OGYAL58TC
44	28	46.7	784	29	CG204197	CG204197 PUFKG68TD
45	28	46.7	807	29	CC667766	CC667766 OGYB049TV

ALIGNMENTS

RESULT 1  
CC821047/c 646 bp DNA linear GSS 22-OCT-2003  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
COMMENT

CC821047  
FECP4B6 Uncultured human fecal virus uncultured human fecal virus  
genomic, genomic survey sequence.

CC821047  
GSS.  
uncultured human fecal virus  
uncultured human fecal virus  
Viruses; environmental samples.

1 (bases 1 to 646)  
Breitbart, M., Hewson, L., Felts, B., Mahafy, J.M., Nulton, J.,  
Salamon, P. and Rohwer, F.  
Metagenomic analyses of an uncultured viral community from human  
feces

J. Bacteriol. 185 (20), 6220-6223 (2003)  
22887472  
14526037  
Contact: Rohwer F  
Biology Dept.  
San Diego State University  
5500 Campanile Dr, San Diego, CA 92102, USA  
Tel: 6195941336  
Fax: 619595676  
Email: forest@sunstroke.sdsu.edu  
Class: shotgun.

FEATURES  
Source  
Location/Qualifiers  
1. 646  
/organism="uncultured human fecal virus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:239364"  
/clone\_lib="Uncultured human fecal virus"

ORIGIN  
Query Match 49.7%; Score 29.8; DB 29; Length 646;  
Best Local Similarity 70.2%; Pred. No. 9.1e+02;  
Matches 40; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 1 GACATCCGCGAGGCCCCCAAGAGCCCTTCGCGACTACGTGACCGCTTCTTCAAG 57  
|||||  
453 GACACCCACACGCGCTCAAGCAGCCCTTATGACTACGTGCTGCCCTTGATCCAG 397  
|||||

RESULT 2 686 bp mRNA linear EST 24-SEP-2003  
LOCUS SCACST3160C06.g ST3 Saccharum officinarum cDNA clone SCACST3160C06  
DEFINITION 5', mRNA sequence.  
CA180823  
CA180823  
CA180823.1 GI:35114987  
KEYWORDS  
SOURCE Saccharum officinarum  
ORGANISM Saccharum officinarum  
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Saccharum.  
1 (bases 1 to 686)  
Vettore,A.L., da Silva,P.R., Kemper,B.L. and Arruda,P.  
The libraries that made SUCEST  
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)  
Contact: Arruda P  
Centro de Biologia Molecular e Engenharia Genetica  
Universidade Estadual de Campinas  
Caixa Postal 6010, 13083-970, Campinas SP, Brazil  
Tel: 55 19 3788 1137  
Fax: 55 19 3788 1089  
Email: paruda@unicamp.br  
Clone distribution: clone distribution information can be found  
through the Brazilian Clone Collection Center (BCCC) at  
http://www.bcccenter.fcav.unesp.br  
Plate: 160 row: C column: 06  
Seq primer: T7 Promoter Primer.  
Location/Qualifiers  
1. 686  
/organism="Saccharum officinarum"  
/mol\_type="mRNA"  
/db\_xref="taxon:4547"  
/clone="SCACST3160C06"  
/lab\_host="DH10B"  
/clone\_lib="ST3"  
/notes="Organ: Fourth apical stalk internodes of adult  
plants; Vector: pSport1; Site 1: SalI; Site 2: NotI; An  
unidirectional cDNA library generated from [Fourth apical  
stalk internodes of adult plants]. cDNA was prepared from  
polyA+ mRNA using Superscript Plasmid System Kit  
(Invitrogen). The double-strand cDNAs were fractionated  
in a sepharose CL-2B 40cm-columns and fragments sizing  
between 0.8 and 1.5 Kb were directionally cloned into the  
vector. Details of each source of RNA and library  
construction can be obtained at  
http://succest.lad.ic.unicamp.br/public"

ORIGIN  
Query Match 49.7%; Score 29.8; DB 14; Length 686;  
Best Local Similarity 70.2%; Pred. No. 9.3e+02;  
Matches 40; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 1 GACATCCGCGAGGCCCCCAAGAGCCCTTCGCGACTACGTGACCGCTTCTTCAAG 57  
|||||  
|||||

Db 11 GCCATGACCCGAGCCTCAAGGCCCGCGTCCGCCGAGCTTGAGAGCTTCTCAAG 67  
|||||

RESULT 3 505 bp DNA linear GSS 28-MAR-2002  
BH786628 fzm014f016d05 fzm filtered library Zea mays genomic clone  
LOCUS fzm014f016d05 5', genomic survey sequence.  
DEFINITION  
BH786628  
BH786628  
BH786628.1 GI:19792775  
KEYWORDS  
SOURCE Zea mays  
ORGANISM Zea mays  
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 505)  
Budiman,M.A., Freese,R.G., Bedell,J.A., Nunberg,A.N. and Lakey,N.D.  
GeneThresher methylation filtered genomic sequences from maize  
Unpublished (2002)  
Contact: Bedell JA  
Orion Genomics, LLC  
4041 Forest Park Ave, St. Louis, MO 63108, USA  
Tel: 314 615 6979  
Fax: 314 615 5975  
Email: jbedell@oriongenomics.com  
Plate: fzm014f016 row: d column: 05  
Seq primer: SK reverse  
Classes: shotgun  
High quality sequence stop: 505.  
Location/Qualifiers  
1. 505  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/cultiivar="Mol17"  
/db\_xref="taxon:4577"  
/clone="fzm014f016d05"  
/clone\_lib="fzm filtered library"  
/notes="Organ: leaf; Vector: pBCK(-); Site 1: HincII; DNA  
prepared from purified nuclei was randomly sheared,  
end-repaired, size fractionated to enrich for the 0.5 to  
5 kb fraction, ligated into HincII-digested pBCK(-)  
vector and electroporated into E. coli cells."

ORIGIN  
Query Match 49.3%; Score 29.6; DB 28; Length 505;  
Best Local Similarity 68.3%; Pred. No. 9.7e+02;  
Matches 41; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 1 GACATCCGCGAGGCCCCCAAGAGCCCTTCGCGACTACGTGACCGCTTCTTCAAG 60  
|||||  
70 GCCCTCTCCCGAGCCACGAGAGCTCTTCTGCGCGTCGACGAGCCCTTCTTCAAG 129  
|||||

RESULT 4 1151 bp mRNA linear EST 06-SEP-2002  
BO619287  
LOCUS RMOSE05H12.SK.ab1 Salt stressed Zea mays roots cDNA library Zea  
DEFINITION RMOSE05H12.SK.ab1 similar to (AF287276) chlorophyll  
a/b-binding protein precursor [Hordeum vulgare], mRNA sequence.  
a/b-binding protein precursor [Hordeum vulgare], mRNA sequence.  
BO619287  
BO619287.1 GI:21621281  
KEYWORDS  
SOURCE Zea mays  
ORGANISM Zea mays  
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 1151)  
Wang,H. and Bohmert,H.J.  
Genomics of plant stress tolerance  
Unpublished (2002)  
Contact: Mark Fredricksen





Seq primer: T7 Promoter Primer.

# FEATURES

source

1. .709  
/organism="Saccharum officinarum"

/mol\_type="mRNA"

/db\_xref="taxon:4547"

/clone="SCCCAM1004D03"

/lab\_host="DH10B"

/clone\_lib="AM1"

/note="Organ: Apical meristem and tissues surrounding of mature plants; Vector: pSPORT1, Site\_1: Salt; Site\_2: NotI; An unidirectional cDNA library generated from [apical meristem and tissues surrounding of mature plants]. cDNA was prepared from polyA+ mRNA using Superscript Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at <http://succest.lad.ic.uticamp.br/public>"

## ORIGIN

Query Match 47.7%; Score 28.6; DB 13; Length 709;  
Best Local Similarity 67.8%; Pred. No. 1.9e+03;  
Matches 40; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 1 GACATCCGCGAGGCCCCCAAGAGCCCTTCGCGACTAGTGACCGCTTTCAGAGC 59  
|||||  
Db 509 GACATCTTCGCGAGCGGAGGCCCCCTTCGCGACTAGTGAGCGGATCCGCAAGCC 567

## RESULT 11

LOCUS BG647389 718 bp mRNA linear EST 24-APR-2001  
DEFINITION EST509008 HOGA Medicago truncatula cDNA clone PHOGA-16B18 5' end,  
mRNA sequence.

ACCESSION BG647389  
VERSION BG647389.1 GI:13782501

KEYWORDS  
SOURCE Medicago truncatula (barrel medic)  
ORGANISM Medicago truncatula

Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae; Medicago.

REFERENCE 1 (bases 1 to 718)  
Hahn, M.G., Ojanen-Reuhs, T., Samac, D., Town, C.D., Van Aken, S.,

TITLE Utterback, T., Cho, J. and Fraser, C.M.  
ESTs from roots of Medicago truncatula treated with

JOURNAL oligogalacturonides of DP 6-20  
COMMENT Unpublished (2001)

CONTACT: Michael G. Hahn  
Complex Carbohydrate Research Center

University of Georgia

220 Riverbend Road, Athens, GA 30602-4712, USA

Tel: 706-542-4457  
Fax: 706-542-4412

Email: hahn@ccrc.uga.edu

G390995e TIGR sequence name: MTMBZ09TK More information is available at: [www.medicago.org](http://www.medicago.org)  
Seq primer: SKmod (CTA GAA CTA GCG GAT CC).

FEATURES  
1. .718  
Location/Qualifiers

/organism="Medicago truncatula"

/mol\_type="mRNA"

/cultivar="A17"

/db\_xref="taxon:3880"

/clone="PHOGA-16B18"

/tissue\_type="3 day old seedling roots"

/dev\_stage="24 hours after treatment in the dark at 26 C with 0.5 mg/ml oligogalacturonides (DP 6-20) in the presence of 100 ug/ml Gentamicin"

/lab\_host="XLOLR"

/clone\_lib="HOGA"

/note="Vector: pBluescript SK-; Site\_1: EcoRI; Site\_2: XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA was directionally ligated into the Unizap XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-Zap phage using Bx-assist helper phage and propagated in SOLR cells."

## ORIGIN

Query Match 47.7%; Score 28.6; DB 12; Length 718;  
Best Local Similarity 67.8%; Pred. No. 1.9e+03;  
Matches 40; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 2 ACATCCGCGAGGCCCCCAAGAGCCCTTCGCGACTAGTGACCGCTTTCAGAGC 60  
|||||  
Db 584 ACCTCCGCGAGAGCCCAAGAGCTCTTCGCGAATACCTCTGATTCAACAGCC 642

## RESULT 12

LOCUS BG549181 525 bp mRNA linear EST 05-APR-2001  
DEFINITION 947073D08.y1 947 - 2 week shoot from Barkan lab Zea mays cDNA, mRNA  
sequence.

ACCESSION BG549181  
VERSION BG549181.1 GI:13557825

KEYWORDS  
SOURCE Zea mays  
ORGANISM Zea mays

Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 525)  
Walbot, V.  
Maize ESTs from various cDNA libraries sequenced at Stanford

JOURNAL University  
COMMENT Unpublished (1999)

CONTACT: Walbot V  
Department of Biological Sciences  
Stanford University  
855 California Ave, Palo Alto, CA 94304, USA

Tel: 650 723 2227  
Fax: 650 725 8221  
Email: walbot@stanford.edu  
Plate: 947073 row: D column: 08.

FEATURES  
1. .525  
Location/Qualifiers

/organism="Zea mays"

/mol\_type="mRNA"

/cultivar="B73"

/db\_xref="taxon:4577"

/tissue\_type="leaf and stem, including leaf base"

/dev\_stage="2 week old seedling (3 leaves)"

/lab\_host="XLI-Blue"

/clone\_lib="947 - 2 week shoot from Barkan lab"

/note="Organ: shoot; Vector: Lambda ZAP (pBluescript SK-); Site\_1: EcoRI; Site\_2: XhoI; Directionally cloned using Stratagene's Unizap XR cDNA cloning kit with the 5' end at the EcoRI site. The library represents 8 x 10e5 independent recombinant phage. The plants were greenhouse grown."

## ORIGIN

Query Match 47.3%; Score 28.4; DB 12; Length 525;  
Best Local Similarity 70.4%; Pred. No. 2e+03;  
Matches 38; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 7 CGCCAGGCGCCCAAGAGCCCTTCGCGACTAGTGACCGCTTTCAGAGC 60  
|||||  
Db 282 CGCAGACATCAAGACCCCGCAGCGTCAACAGACCCCATCTCAAGAGC 335

## RESULT 13



CA210513 544 bp mRNA linear EST 25-SEP-2003  
 LOCUS SCEPSB1128C11.9 SBI Saccharum officinarum cDNA clone SCEPSB1128C11  
 DEFINITION 5', mRNA sequence.  
 CA210513  
 ACCESSION CA210513.1 GI:35253896  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Saccharum officinarum  
 Saccharum officinarum  
 Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoideae; Andropogoneae; Saccharum.  
 1 (bases 1 to 544)  
 Vettore,A.L., da Silva,P.R., Kemper,B.L. and Arruda,P.  
 The libraries that made SUCEST  
 Genet. Mol. Biol. 24 (1-4), 1-7 (2001)  
 COMMENT  
 Contact: Arruda P  
 Centro de Biologia Molecular e Engenharia Genetica  
 Universidade Estadual de Campinas  
 Caixa Postal 6010, 13083-970, Campinas SP, Brazil  
 Tel: 55 19 3788 1137  
 Fax: 55 19 3788 1089  
 Email: parrruda@unicamp.br  
 Clone distribution information can be found  
 through the Brazilian Clone Collection Center (BCCC) at  
 http://www.bcccenter.fcav.unesp.br  
 Plate: 128 row: C column: 11  
 Seq primer: T7 Promoter Primer.  
 Location/Qualifiers  
 1..544  
 /organism="Saccharum officinarum"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4547"  
 /clone="SCEPSB1128C11"  
 /lab\_host="DH10B"  
 /clone\_lib="SBI"  
 /note="Organ: Stalk Bark from adult plants; Vector:  
 pSport1; Site\_1: SalI; Site\_2: NotI; An unidirectional  
 cDNA library generated from [Stalk Bark from adult  
 plants]. cDNA was prepared from polyA+ mRNA using  
 Superscript Plasmid System Kit (Invitrogen). The  
 double-strand cDNAs were fractionated in a sepharose  
 CL-2B 40cm-columns and fragments sizing between 0.8 and  
 1.5 Kb were directionally cloned into the vector. Details  
 of each source of RNA and library construction can be  
 obtained at http://sucest.lad.ic.unicamp.br/public"

ORIGIN  
 Query Match 47.3%; Score 28.4; DB 14; Length 544;  
 Best Local Similarity 70.4%; Pred. No. 2e+03; Mismatches 16; Indels 0; Gaps 0;  
 Matches 38; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

OR  
 1 GACATCGCGCAGGGCCCCAAGAGAGCCCTTCGCGACTACGTGACCGCTTCTC 54  
 |||||  
 254 GACTTCGCGCGAGCCTCGAGAGACTCGTCTGCAACCACTCGGCGGCTGCTTC 307  
 |||||

RESULT 14  
 CG250176 577 bp DNA linear GSS 25-AUG-2003  
 LOCUS OGVCW47TV ZM 0.7-1.5 KB Zea mays genomic clone ZMMBMA0507H21,  
 DEFINITION genomic survey sequence.  
 CG250176  
 ACCESSION CG250176.1 GI:34152266  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Zea mays  
 Zea mays  
 Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoideae; Andropogoneae; Zea.  
 1 (bases 1 to 577)  
 Whiteclaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,  
 Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,

TITLE  
 JOURNAL  
 COMMENT  
 Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.  
 Consortium for Maize Genomics  
 Unpublished (2002)  
 Contact: Cathy Whiteclaw  
 TIGR  
 9712 Medical Center Drive, Rockville, MD 20850, USA  
 Tel: 301-838-5843  
 Fax: 301-838-0208  
 Email: whiteclaw@tigr.org  
 Seg primer: TF  
 Clae: sheared ends.  
 Location/Qualifiers  
 1..577  
 /organism="Zea mays"  
 /mol\_type="genomic DNA"  
 /strain="B73"  
 /db\_xref="taxon:4577"  
 /clone="ZMMBMA0507H21"  
 /clone\_lib="ZM 0.7-1.5 KB"  
 /note="Vector: pBCK-; Site\_1: HincII; 0.7-1.5 kb  
 methylation filtered genomic DNA library"

ORIGIN  
 Query Match 47.3%; Score 28.4; DB 29; Length 577;  
 Best Local Similarity 70.4%; Pred. No. 2e+03; Mismatches 16; Indels 0; Gaps 0;  
 Matches 38; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

OR  
 1 GACATCGCGCAGGGCCCCAAGAGAGCCCTTCGCGACTACGTGACCGCTTCTC 54  
 |||||  
 378 GACTTCGCGCGAGCCTCGAGAGACTCGTCTGCAACCACTCGGCGGCTGCTTC 431  
 |||||

RESULT 15  
 BG517389 580 bp mRNA linear EST 30-MAR-2001  
 LOCUS 947061G10.y1 947 - 2 week shoot from Barkan lab Zea mays cDNA, mRNA  
 DEFINITION sequence.  
 BG517389  
 ACCESSION BG517389.1 GI:13490625  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Zea mays  
 Zea mays  
 Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoideae; Andropogoneae; Zea.  
 1 (bases 1 to 580)  
 Walbot,V.  
 Maize ESTs from various cDNA libraries sequenced at Stanford  
 University  
 Unpublished (1999)  
 Contact: Walbot V  
 Department of Biological Sciences  
 Stanford University  
 855 California Ave, Palo Alto, CA 94304, USA  
 Tel: 650 723 2227  
 Fax: 650 725 8221  
 Email: walbot@stanford.edu  
 Plate: 947061 row: G column: 10.  
 Location/Qualifiers  
 1..580  
 /organism="Zea mays"  
 /mol\_type="mRNA"  
 /cultivar="B73"  
 /db\_xref="taxon:4577"  
 /tissue\_type="leaf and stem, including leaf base"  
 /dev\_stage="2 week old seedling (3 leaves)"  
 /lab\_host="XL1-Blue"  
 /clone\_lib="947 - 2 week shoot from Barkan lab"  
 /note="Organ: shoot; Vector: Lambda ZAP (pBluescript SK-);  
 Site\_1: EcoRI; Site\_2: XhoI; Directionally cloned using  
 Strategene's Unizap cDNA cloning kit with the 5' end  
 at the EcoRI site. The library represents 8 x 10e5  
 independent recombinant phage. The plants were greenhouse

FEATURES  
 source





APPLICANT: GREER, Catherine  
APPLICANT: SELBY, Mark  
APPLICANT: WALKER, Christopher  
TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION  
TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES  
FILE REFERENCE: 1621.002  
CURRENT APPLICATION NUMBER: US/09/475,515A  
CURRENT FILING DATE: 1999-12-30  
NUMBER OF SEQ ID NOS: 90  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 9  
LENGTH: 1268  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic Gag  
US-09-475-515-9

Query Match 89.3%; Score 53.6; DB 4; Length 1268;  
Best Local Similarity 93.3%; Pred. No. 7.8e-09;  
Matches 56; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 GACATCAAGCAGGCGCCCAAGAGCCCTTCGCGACTACGTGAGCCGCTTCTTCAAGACC 60  
DB 862 GACATCCGCGCAGGCGCCCAAGAGCCCTTCGCGACTACGTGAGCCGCTTCTTCAAGACC 921

## RESULT 3

US-09-475-515-4  
Sequence 4, Application US/09475515A  
Patent No. 6602705  
GENERAL INFORMATION:  
APPLICANT: BARNETT, Susan  
APPLICANT: ZUR MEGEDB, Jan  
APPLICANT: SRIVASTAVA, Indresh  
APPLICANT: LIAN, Ying  
APPLICANT: HARTOG, Karin  
APPLICANT: LIU, Hong  
APPLICANT: GREER, Catherine  
APPLICANT: SELBY, Mark  
APPLICANT: WALKER, Christopher  
TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION  
TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES  
FILE REFERENCE: 1621.002  
CURRENT APPLICATION NUMBER: US/09/475,515A  
CURRENT FILING DATE: 1999-12-30  
NUMBER OF SEQ ID NOS: 90  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 4  
LENGTH: 1515  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-09-475-515-4

Query Match 89.3%; Score 53.6; DB 4; Length 1515;  
Best Local Similarity 93.3%; Pred. No. 7.9e-09;  
Matches 56; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 GACATCAAGCAGGCGCCCAAGAGCCCTTCGCGACTACGTGAGCCGCTTCTTCAAGACC 60  
DB 862 GACATCCGCGCAGGCGCCCAAGAGCCCTTCGCGACTACGTGAGCCGCTTCTTCAAGACC 921

## RESULT 4

US-09-475-515-5  
Sequence 5, Application US/09475515A  
Patent No. 6602705  
GENERAL INFORMATION:  
APPLICANT: BARNETT, Susan

APPLICANT: ZUR MEGEDB, Jan  
APPLICANT: SRIVASTAVA, Indresh  
APPLICANT: LIAN, Ying  
APPLICANT: HARTOG, Karin  
APPLICANT: LIU, Hong  
APPLICANT: GREER, Catherine  
APPLICANT: SELBY, Mark  
APPLICANT: WALKER, Christopher  
TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION  
TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES  
FILE REFERENCE: 1621.002  
CURRENT APPLICATION NUMBER: US/09/475,515A  
CURRENT FILING DATE: 1999-12-30  
NUMBER OF SEQ ID NOS: 90  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 5  
LENGTH: 1853  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-09-475-515-5

Query Match 89.3%; Score 53.6; DB 4; Length 1853;  
Best Local Similarity 93.3%; Pred. No. 8.1e-09;  
Matches 56; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 GACATCAAGCAGGCGCCCAAGAGCCCTTCGCGACTACGTGAGCCGCTTCTTCAAGACC 60  
DB 862 GACATCCGCGCAGGCGCCCAAGAGCCCTTCGCGACTACGTGAGCCGCTTCTTCAAGACC 921

## RESULT 5

US-09-475-515-78  
Sequence 78, Application US/09475515A  
Patent No. 6602705  
GENERAL INFORMATION:  
APPLICANT: BARNETT, Susan  
APPLICANT: ZUR MEGEDB, Jan  
APPLICANT: SRIVASTAVA, Indresh  
APPLICANT: LIAN, Ying  
APPLICANT: HARTOG, Karin  
APPLICANT: LIU, Hong  
APPLICANT: GREER, Catherine  
APPLICANT: SELBY, Mark  
APPLICANT: WALKER, Christopher  
TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION  
TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES  
FILE REFERENCE: 1621.002  
CURRENT APPLICATION NUMBER: US/09/475,515A  
CURRENT FILING DATE: 1999-12-30  
NUMBER OF SEQ ID NOS: 90  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 78  
LENGTH: 1865  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: GPI  
US-09-475-515-78

Query Match 89.3%; Score 53.6; DB 4; Length 1865;  
Best Local Similarity 93.3%; Pred. No. 8.1e-09;  
Matches 56; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 GACATCAAGCAGGCGCCCAAGAGCCCTTCGCGACTACGTGAGCCGCTTCTTCAAGACC 60  
DB 868 GACATCCGCGCAGGCGCCCAAGAGCCCTTCGCGACTACGTGAGCCGCTTCTTCAAGACC 927

## RESULT 6

US-09-475-515-79

```
; Sequence 79, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; FILE REFERENCE: 1621.002
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 79
; LENGTH: 1865
; TYPE: DNA
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Description of Artificial Sequence: GP2
US-09-475-515-79
```

```
Query Match      89.3%; Score 53.6; DB 4; Length 1865;
Best Local Similarity 93.3%; Pred. No. 8.1e-09;
Matches 56; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
OY      1 GACATCAAGCAGGCGCCCAAGAGAGCCCTTCGCGACTACGTGACCGCTTCTCAAGACC 60
Db      868 GACATCCGCGCAGGCGCCCAAGAGAGCCCTTCGCGACTACGTGACCGCTTCTCAAGACC 927
```

## RESULT 7

```
US-09-475-515-7
; Sequence 7, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 2031
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: HIV-Gag/HCV-core fusion polypeptide
US-09-475-515-7
```

```
Query Match      89.3%; Score 53.6; DB 4; Length 2031;
Best Local Similarity 93.3%; Pred. No. 8.1e-09;
Matches 56; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
OY      1 GACATCAAGCAGGCGCCCAAGAGAGCCCTTCGCGACTACGTGACCGCTTCTCAAGACC 60
Db      862 GACATCCGCGCAGGCGCCCAAGAGAGCCCTTCGCGACTACGTGACCGCTTCTCAAGACC 921
```

```
RESULT 8
US-09-475-515-6
; Sequence 6, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
```

```
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 4319
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-475-515-6
```

```
Query Match      89.3%; Score 53.6; DB 4; Length 4319;
Best Local Similarity 93.3%; Pred. No. 8.8e-09;
Matches 56; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
OY      1 GACATCAAGCAGGCGCCCAAGAGAGCCCTTCGCGACTACGTGACCGCTTCTCAAGACC 60
Db      862 GACATCCGCGCAGGCGCCCAAGAGAGCCCTTCGCGACTACGTGACCGCTTCTCAAGACC 921
```

## RESULT 9

```
US-09-475-515-75
; Sequence 75, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 75
; LENGTH: 4472
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
US-09-475-515-75
```

```
Query Match      89.3%; Score 53.6; DB 4; Length 4472;
Best Local Similarity 93.3%; Pred. No. 8.8e-09;
```



Matches 56; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GACATCAAGCAGGGCCCCCAAGAGCCCTTCGCGACTAGTGAGACCGCTTCTTCAAGACC 60  
DB 3794 GACATCCGCGCAGGGCCCCCAAGAGCCCTTCGCGACTAGTGAGACCGCTTCTTCAAGACC 3853

## RESULT 10

US-09-475-515-76  
Sequence 76, Application US/09475515A

Patent No. 6602705

GENERAL INFORMATION:

APPLICANT: BARNETT, Susan

APPLICANT: ZUR MEGEDE, Jan

APPLICANT: SRIVASTAVA, Indresh

APPLICANT: LIAN, Ying

APPLICANT: HARTOG, Karin

APPLICANT: LIU, Hong

APPLICANT: GREER, Catherine

APPLICANT: SELBY, Mark

APPLICANT: WALKER, Christopher

TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION

TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES

FILE REFERENCE: 1621.002

CURRENT APPLICATION NUMBER: US/09/475, 515A

CURRENT FILING DATE: 1999-12-30

NUMBER OF SEQ ID NOS: 90

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 76

LENGTH: 4608

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURES:

OTHER INFORMATION: Description of Artificial Sequence:

US-09-475-515-76

Query Match 89.3%; Score 53.6; DB 4; Length 4608;  
Best Local Similarity 93.3%; Pred. No. 8.8e-09;  
Matches 56; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GACATCAAGCAGGGCCCCCAAGAGCCCTTCGCGACTAGTGAGACCGCTTCTTCAAGACC 60  
DB 3930 GACATCCGCGCAGGGCCCCCAAGAGCCCTTCGCGACTAGTGAGACCGCTTCTTCAAGACC 3989

## RESULT 11

US-09-475-515-74

Sequence 74, Application US/09475515A

Patent No. 6602705

GENERAL INFORMATION:

APPLICANT: BARNETT, Susan

APPLICANT: ZUR MEGEDE, Jan

APPLICANT: SRIVASTAVA, Indresh

APPLICANT: LIAN, Ying

APPLICANT: HARTOG, Karin

APPLICANT: LIU, Hong

APPLICANT: GREER, Catherine

APPLICANT: SELBY, Mark

APPLICANT: WALKER, Christopher

TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION

TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES

FILE REFERENCE: 1621.002

CURRENT APPLICATION NUMBER: US/09/475, 515A

CURRENT FILING DATE: 1999-12-30

NUMBER OF SEQ ID NOS: 90

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 74

LENGTH: 4689

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURES:

OTHER INFORMATION: Description of Artificial Sequence:

OTHER INFORMATION: gp160.modSF162.gag.modSF2  
US-09-475-515-74

Query Match 89.3%; Score 53.6; DB 4; Length 4689;  
Best Local Similarity 93.3%; Pred. No. 8.9e-09;  
Matches 56; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GACATCAAGCAGGGCCCCCAAGAGCCCTTCGCGACTAGTGAGACCGCTTCTTCAAGACC 60  
DB 4011 GACATCCGCGCAGGGCCCCCAAGAGCCCTTCGCGACTAGTGAGACCGCTTCTTCAAGACC 4070

## RESULT 12

US-09-475-515-73

Sequence 73, Application US/09475515A

Patent No. 6602705

GENERAL INFORMATION:

APPLICANT: BARNETT, Susan

APPLICANT: ZUR MEGEDE, Jan

APPLICANT: SRIVASTAVA, Indresh

APPLICANT: LIAN, Ying

APPLICANT: HARTOG, Karin

APPLICANT: LIU, Hong

APPLICANT: GREER, Catherine

APPLICANT: SELBY, Mark

APPLICANT: WALKER, Christopher

TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION

TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES

FILE REFERENCE: 1621.002

CURRENT APPLICATION NUMBER: US/09/475, 515A

CURRENT FILING DATE: 1999-12-30

NUMBER OF SEQ ID NOS: 90

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 73

LENGTH: 4766

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURES:

OTHER INFORMATION: Description of Artificial Sequence:

US-09-475-515-73

Query Match 89.3%; Score 53.6; DB 4; Length 4766;  
Best Local Similarity 93.3%; Pred. No. 8.9e-09;  
Matches 56; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GACATCAAGCAGGGCCCCCAAGAGCCCTTCGCGACTAGTGAGACCGCTTCTTCAAGACC 60  
DB 4088 GACATCCGCGCAGGGCCCCCAAGAGCCCTTCGCGACTAGTGAGACCGCTTCTTCAAGACC 4147

## RESULT 13

US-09-552-950-2

Sequence 2, Application US/09552950

Patent No. 6541248

GENERAL INFORMATION:

APPLICANT: Oxford Biomedica (UK) Limited

APPLICANT: Anti-Viral Vectors

FILE REFERENCE: 674524-2004

CURRENT APPLICATION NUMBER: US/09/552, 950

CURRENT FILING DATE: 2000-04-20

NUMBER OF SEQ ID NOS: 22

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 2

LENGTH: 4307

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURES:

OTHER INFORMATION: Description of Artificial Sequence:gagpol-SynGP - codon

US-09-552-950-2

Query Match 71.7%; Score 43; DB 4; Length 4307;



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 28, 2004, 08:01:50 ; Search time 234.634 Seconds  
(without alignments)  
7636.275 Million cell updates/sec

Title: US-09-475-704A-1

Perfect score: 60  
Sequence: 1 gacatcaagcagggcccca.....tgagccgtcttccaagacc 60

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST: \*  
1: em\_estba: \*  
2: em\_esthum: \*  
3: em\_estin: \*  
4: em\_estmu: \*  
5: em\_estov: \*  
6: em\_estpl: \*  
7: em\_estro: \*  
8: em\_htc: \*  
9: gb\_est1: \*  
10: gb\_est2: \*  
11: gb\_htc: \*  
12: gb\_est3: \*  
13: gb\_est4: \*  
14: gb\_est5: \*  
15: em\_estfun: \*  
16: em\_estom: \*  
17: em\_gss\_hum: \*  
18: em\_gss\_iny: \*  
19: em\_gss\_pln: \*  
20: em\_gss\_vrt: \*  
21: em\_gss\_fun: \*  
22: em\_gss\_mam: \*  
23: em\_gss\_mus: \*  
24: em\_gss\_pro: \*  
25: em\_gss\_rtd: \*  
26: em\_gss\_pbg: \*  
27: em\_gss\_vrl: \*  
28: gb\_gss1: \*  
29: gb\_gss2: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	29.8	49.7	686	14	CA180823 SCACST316
2	29.6	49.3	568	6	Al809870 Triticum
3	29.6	49.3	1151	13	BQ619287 RNOSQ5H1
4	29.6	49.3	1151	13	BQ619288 RNOSQ6A0

5	29.6	49.3	1151	13	BQ619295	BQ619295 RNOSQ6A0
6	29.6	49.3	1151	13	BQ619376	BQ619376 RNOSQ7A0
7	29.6	49.3	1151	13	BQ619455	BQ619455 RNOSQ7H0
8	29.6	49.3	1151	13	BQ619551	BQ619551 C11aetPC
9	29.6	49.3	1161	14	CK166030	CK166030 FGAS0507
10	29.2	48.7	708	13	BQ645917	BU645917 1112034B0
11	29.2	48.7	922	14	CK165296	CK165296 FGAS04924
12	29.2	48.3	781	14	CP430199	CP430199 PH1_26_D0
13	28.6	47.7	455	6	AL811116	AL811116 Triticum
14	28.6	47.7	491	9	AU031268	AU031268 AU031268
15	28.6	47.7	513	14	CA707887	CA707887 wdk2c.pk0
16	28.6	47.7	582	6	AL811117	AL811117 Triticum
17	28.6	47.7	875	14	CK164904	CK164904 FGAS04883
18	28.4	47.3	560	12	BG517417	BG517417 947062C12
19	28.2	47.0	408	14	CB639810	CB639810 OSJNEA12B
20	28.2	47.0	576	29	CG225792	CG225792 OG0CF27TH
21	28.2	47.0	576	29	CG225799	CG225799 OG0CF27TV
22	28.2	47.0	585	10	AM147097	AM147097 707011F07
23	28.2	47.0	597	29	CG675574	CG675574 OGJAX28TV
24	28.2	47.0	625	14	CD884980	CD884980 F1_118H11
25	28.2	47.0	646	29	CC821047	CC821047 FBCPA86 U
26	28.2	47.0	675	29	CG249691	CG249691 OEXCO69TH
27	28.2	47.0	809	14	CK198615	CK198615 FGAS00710
28	28.2	47.0	819	29	CG293655	CG293655 OG0B73TH
29	28.2	47.0	819	29	CG349088	CG349088 OGYBL67TV
30	28.2	47.0	823	29	CG288932	CG288932 OGTC45TV
31	28.2	47.0	829	29	CC682396	CC682396 OGLBB77TV
32	28.2	47.0	829	29	CG313210	CG313210 OGXDR22TV
33	28.2	47.0	858	29	CG364851	CG364851 OG1AE85TV
34	28.2	47.0	869	29	CC707427	CC707427 OGUBB52TV
35	28.2	47.0	923	29	CC731054	CC731054 OGVBI45TV
36	28.2	46.7	120	14	CD955986	CD955986 SBY_233 G
37	28.2	46.7	168	10	BG050418	BG050418 FM1_53_H0
38	28.2	46.7	216	9	AA143927	AA143927 zEST00762
39	28.2	46.7	219	10	BR356040	BR356040 DG1_121_H
40	28.2	46.7	262	14	CF061008	CF061008 OCT16d08
41	28.2	46.7	282	12	BI992820	BI992820 1020068G0
42	28.2	46.7	305	13	BQ657317	BQ657317 HA08N16u
43	28.2	46.7	306	12	BJ240990	BJ240990 BU240990
44	28.2	46.7	327	10	BF317794	BF317794 OVI_9_E10
45	28.2	46.7	328	14	CA999329	CA999329 S234G_C07

#### ALIGNMENTS

RESULT 1  
LOCUS CA180823 686 bp mRNA linear EST 24-SEP-2003  
DEFINITION SCACST316OC06.g ST3 Saccharum officinarum cDNA clone SCACST316OC06  
5', mRNA sequence.  
ACCESSION CA180823  
VERSION CA180823.1 GI:35114987  
KEYWORDS EST.  
SOURCE Saccharum officinarum  
ORGANISM Saccharum officinarum  
REFERENCE 1 (bases 1 to 686)  
Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.  
The libraries that made SUCEST  
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)  
CONTACT: Arruda P  
Centro de Biologia Molecular e Engenharia Genetica  
Universidade Estadual de Campinas  
Caixa Postal 6010, 13083-970, Campinas SP, Brazil  
Tel: 55 19 3788 1137  
Fax: 55 19 3788 1089  
Email: parruda@unicamp.br  
Clone distribution: clone distribution information can be found  
through the Brazilian Clone Collection Center (BCCC) at  
http://www.bcccenter.fcav.unesp.br



```

SOURCE      Zea mays
ORGANISM    Zea mays
REFERENCE   Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS     Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
TITLE       clade; Panicoideae; Andropogoneae; Zea.
JOURNAL
COMMENT     1 (bases 1 to 1151)
            Wang, H. and Bohnert, H.J.
            Genomics of plant stress tolerance
            Unpublished (2002)
            Contact: Mark Fredricksen
            Department of Plant Biology
            University of Illinois
            1201 W. Gregory Dr., Urbana, IL 61801, USA
            Tel: 2172655473
            Email: bohnertlab@life.uiuc.edu.

FEATURES
  source
    1..1151
      /organism="Zea mays"
      /mol_type="mRNA"
      /cultivar="B73"
      /db_xref="taxon:4577"
      /clone="RNOSBQ6A01.SK.ab1"
      /tissue_type="Roots"
      /dev_stage="2 weeks old"
      /clone_lib="Salt stressed Zea mays roots cDNA library"
      /note="Vector: pBluescript SK+; Stressed 24 hours at 150
mm NaCl"

ORIGIN

Query Match          49.3%; Score 29.6; DB 13; Length 1151;
Best Local Similarity 73.1%; Pred. No. 7.1e+02;
Matches 38; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Db      524 GCAGGACATCAAGAACCCCGTCAGCGTCAACCAAGACCCCATCTTCAAGAGC 575
Oy      9 GCAGGCCCCCAAGAGACCCCTTCCGCACTACGTGACCGCTTTCAGAGCC 60
        ||||| | ||||| ||||| ||||| ||||| ||||| ||||| |||||
LOCUS     BQ619295                      1151 bp      mRNA      linear      EST 06-SEP-2002
DEFINITION RNOSBQ6A08.SK.ab1 Salt stressed Zea mays roots cDNA library Zea
            mays cDNA clone RNOSBQ6A08.SK.ab1 similar to (AF287276) chlorophyll
            a/b-binding protein precursor [Hordeum vulgare], mRNA sequence.
ACCESSION BQ619295
VERSION   BQ619295.1 GI:21621289
KEYWORDS  EST.
SOURCE    Zea mays
ORGANISM  Zea mays
REFERENCE Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS     Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
TITLE       clade; Panicoideae; Andropogoneae; Zea.
JOURNAL
COMMENT     1 (bases 1 to 1151)
            Wang, H. and Bohnert, H.J.
            Genomics of plant stress tolerance
            Unpublished (2002)
            Contact: Mark Fredricksen
            Department of Plant Biology
            University of Illinois
            1201 W. Gregory Dr., Urbana, IL 61801, USA
            Tel: 2172655473
            Email: bohnertlab@life.uiuc.edu.

FEATURES
  source
    1..1151
      /organism="Zea mays"
      /mol_type="mRNA"
      /cultivar="B73"
      /db_xref="taxon:4577"
      /clone="RNOSBQ6A08.SK.ab1"
      /tissue_type="Roots"
      /dev_stage="2 weeks old"
      /clone_lib="Salt stressed Zea mays roots cDNA library"
      /note="Vector: pBluescript SK+; Stressed 24 hours at 150

```

ORIGIN	mm NaCl"
Query Match	49.3%; Score 29.6; DB 13; Length 1151;
Best Local Similarity	73.1%; Pred. No. 7.1e+02;
Matches	38; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
Db	524 GCAGGACATCAAGAACCCCGCTCAGCGTCAACCAAGACCCCATCTTCAAGAGC 575
Oy	9 GCAGGAGCCCAAGAGAGCCCTTCCGCGACTACGTGAGCCGCTTCTTCAAGACC 60
LOCUS	1151 bp mRNA linear EST 06-SEP-2002
DEFINITION	RNASEQ07A07.SK.ab1 Salt stressed Zea mays roots cDNA library Zea
ACCESSION	BO619376
VERSION	BO619376.1 GI:21621370
KEYWORDS	EST.
SOURCE	Zea mays
ORGANISM	Zea mays
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
TITLE	ciade; Panicoidae; Andropogoneae; Zea.
JOURNAL	1 (bases 1 to 1151)
COMMENT	Wang, H. and Bohnert, H.J. Genomics of plant stress tolerance Unpublished (2002) Contact: Mark Fredricksen Department of Plant Biology University of Illinois 1201 W. Gregory Dr., Urbana, IL 61801, USA Tel: 2172655473 Email: bohnertlab@life.uiuc.edu.
FEATURES	Location/Qualifiers
source	1..1151
	/organism="Zea mays"
	/mol_type="mRNA"
	/culivar="B73"
	/db_xref="taxon:4577"
	/clone="RNASEQ07A07.SK.ab1"
	/tissue_type="Roots"
	/dev_stage="2 weeks old"
	/clone_lib="Salt stressed Zea mays roots cDNA library"
	/note="Vector: pBluescript SK+; Stressed 24 hours at 150
	mM NaCl"
ORIGIN	
Query Match	49.3%; Score 29.6; DB 13; Length 1151;
Best Local Similarity	73.1%; Pred. No. 7.1e+02;
Matches	38; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
Oy	9 GCAGGAGCCCAAGAGAGCCCTTCCGCGACTACGTGAGCCGCTTCTTCAAGACC 60
Db	524 GCAGGACATCAAGAACCCCGCTCAGCGTCAACCAAGACCCCATCTTCAAGAGC 575
RESULT 7	
LOCUS	1151 bp mRNA linear EST 06-SEP-2002
DEFINITION	RNASEQ07H05.SK.ab1 Salt stressed Zea mays roots cDNA library Zea
ACCESSION	BO619455
VERSION	BO619455.1 GI:21621449
KEYWORDS	EST.
SOURCE	Zea mays
ORGANISM	Zea mays
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
	ciade; Panicoidae; Andropogoneae; Zea.



REFERENCE 1 (bases 1 to 1151)  
 AUTHORS Wang, H. and Bohnert, H.J.  
 TITLE Genomics of plant stress tolerance  
 JOURNAL Unpublished (2002)  
 COMMENT Contact: Mark Fredrickson  
 Department of Plant Biology  
 University of Illinois  
 1201 W. Gregory Dr., Urbana, IL 61801, USA  
 Tel: 2172655473  
 Email: bohnertlab@life.uiuc.edu.

FEATURES  
 source Location/Qualifiers  
 1..1151  
 /organism="Zea mays"  
 /mol\_type="mRNA"  
 /cultivar="B73"  
 /db\_xref="taxon:4577"  
 /clone="RNOSEQ7H05\_SK.ab1"  
 /issue\_type="Roots"  
 /dev\_stage="2 weeks old"  
 /clone\_lib="Salt stressed Zea mays roots cDNA library"  
 /note="Vector: pBluescript SK+; Stressed 24 hours at 150 mM NaCl"

## ORIGIN

Query Match 49.3%; Score 29.6; DB 13; Length 1151;  
 Best Local Similarity 73.1%; Pred. No. 7.1e+02;  
 Matches 38; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 9 GCAGGCCCCAGAGAGCCCTTCGCGACTACGTCGACCGCTTCTCAAGACC 60  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 524 GCAGACATCAAGAACCCCGTCAGCGTCAACCAAGACCCCATCTTCAAGAGC 575

RESULT 8  
 LOCUS B0619551 1151 bp mRNA linear EST 06-SEP-2002  
 DEFINITION C11a1stPCR SK.ab1 Salt stressed Zea mays leaves cDNA library Zea  
 mays cDNA clone C11a1stPCR SK.ab1 similar to (AF287276)  
 chlorophyll a/b-binding protein precursor [Hordeum vulgare], mRNA  
 sequence.  
 B0619551  
 B0619551.1 GI:21621545  
 EST.

ACCESSION B0619551  
 VERSION B0619551.1  
 KEYWORDS EST.  
 SOURCE Zea mays  
 ORGANISM Zea mays  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 1151)  
 AUTHORS Wang, H. and Bohnert, H.J.  
 TITLE Genomics of plant stress tolerance  
 JOURNAL Unpublished (2002)  
 COMMENT Contact: Mark Fredrickson  
 Department of Plant Biology  
 University of Illinois  
 1201 W. Gregory Dr., Urbana, IL 61801, USA  
 Tel: 2172655473  
 Email: bohnertlab@life.uiuc.edu.

FEATURES  
 source Location/Qualifiers  
 1..1151  
 /organism="Zea mays"  
 /mol\_type="mRNA"  
 /cultivar="B73"  
 /db\_xref="taxon:4577"  
 /clone="C11a1stPCR SK.ab1"  
 /issue\_type="Leaves"  
 /dev\_stage="2 weeks old"  
 /clone\_lib="Salt stressed Zea mays leaves cDNA library"  
 /note="Vector: pBluescript SK+; Stressed 24 hours at 150 mM NaCl"

## ORIGIN

Query Match 49.3%; Score 29.6; DB 13; Length 1151;

Best Local Similarity 73.1%; Pred. No. 7.1e+02;  
 Matches 38; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 9 GCAGGCCCCAGAGAGCCCTTCGCGACTACGTCGACCGCTTCTCAAGACC 60  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 524 GCAGACATCAAGAACCCCGTCAGCGTCAACCAAGACCCCATCTTCAAGAGC 575

RESULT 9  
 LOCUS CK166030 1161 bp mRNA linear EST 05-DEC-2003  
 DEFINITION FGAS050078 Triticum aestivum FGAS: Talt7 Triticum aestivum cDNA,  
 mRNA sequence.  
 CK166030  
 CK166030.1 GI:38998675  
 EST.

ACCESSION CK166030  
 VERSION CK166030  
 KEYWORDS EST.  
 SOURCE Triticum aestivum (bread wheat)  
 ORGANISM Triticum aestivum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Poideae; Triticeae; Triticum.

REFERENCE 1 (bases 1 to 1161)  
 AUTHORS Allard, F., Crosby, W.L., Danyluk, J., Eudes, F., Frick, M., Gaudet, D.,  
 Gensewein, B., Graf, R., Gulick, P., Hrycan, L.D., Laroche, A.,  
 Links, M.G., McCarthy, E.L., Monroy, A., Muzak, I., Nilson, D.,  
 Peniket, C., Roach, J.L. and Sarhan, F.  
 Functional Genomics of Abiotic Stress in Wheat and Canola Crops  
 Unpublished (2003)

COMMENT Contact: Mm L Crosby  
 Bioinformatics  
 University of Saskatchewan, Department of Computer Science  
 1C101 Engineering Building, 57 Campus Drive, Saskatoon,  
 Saskatchewan, S7N 5A9, Canada  
 Tel: 306 966 1769  
 Fax: 306 966 2033  
 Email: fgas\_estseqs.usask.ca

This sequence is the direct result of the Base calling software  
 phred (default parameters). It is the raw base calls. To aid in the  
 identification of the high quality insert the software Lucy  
 (default parameters) has been run on this sequence. Lucy identified  
 the region [62,774].  
 Plate: Talt707 row: H column: 16.

FEATURES  
 source Location/Qualifiers  
 1..1161  
 /organism="Triticum aestivum"  
 /mol\_type="mRNA"  
 /cultivar="Wheat line CI 14106"  
 /db\_xref="taxon:4565"  
 /lab\_host="DH5 alpha"  
 /clone\_lib="Triticum aestivum FGAS: Talt7"  
 /note="Organ: Crown; Vector: pGEM-T; SSH (suppression  
 subtractive hybridization) cDNA library from genotype  
 CI14106 cold hardened at 2 C for 21 days and 49 days  
 (equal amount of cDNA pooled together before subtraction,  
 tester) and subtracted against genotype CI14106  
 non-hardened (20 C) (driver). Nitro-pyrole anchored  
 oligo-dT priming and non-directional cloning."

## ORIGIN

Query Match 49.3%; Score 29.6; DB 14; Length 1161;  
 Best Local Similarity 68.3%; Pred. No. 7.1e+02;  
 Matches 41; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 1 GACATCAAGCAGGCCCCAAGAGCCCTTCGCGACTACGTGAGACCGCTTCTCAAGACC 60  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 554 GTCATCAAGAGACCGCGCCCAAGACCTTCACGACCTCGAGCAGCACCTCCTCAGAGAC 613

RESULT 10  
 LOCUS BU645917 708 bp mRNA linear EST 30-SEP-2002  
 DEFINITION 1112034B07.x1 C. reinhardtii CC-1690 (mt+), CC-1691 (mt-), Gamete  
 (normalized), Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA

```
sequence.
BU645917
BU645917.1 GI:23358097
EST.
Chlamydomonas reinhardtii
Chlamydomonas reinhardtii
Bukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.
REFERENCE
AUTHORS Grossman,A., Chang,C.-W., Davies,J., Harris,B., Hauser,C.,
Lefebvre,P., McDermott,J.P., Shrager,J., Silflow,C. and Stern,D.
TITLE Analyses of the Chlamydomonas reinhardtii Genome: A Model,
Unicellular System for Analyzing Gene Function and Regulation in
Vascular Plants. Project: 1112
JOURNAL Unpublished (2002)
COMMENT Contact: Charles Hauser
DCMB Box 91000
Duke University
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu.
FEATURES
Source
Location/Qualifiers
1..708
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="21gr (CC-1690 wild type mt+) & 6145c (CC-1691
wild type mt-)"
/db_xref="taxon:3055"
/cclone_lib="C. reinhardtii CC-1690 (mt+), CC-1691 (mt-),
Gamete (normalized), Lambda Zap II"
/note="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2:
XhoI; Gamete library was constructed by Hui Zhao, Min Lu,
Jeffrey McDermott, William J. Snell and John Davies.
Strain 21gr cells (CC-1690; mating type plus) and strain
6145c cells(CC-1691; mating type minus) that had been
growing on a light-dark cycle (13:11 L/D) in R-medium
(Sager and Granick) were separately transferred into
nitrogen-free medium at 8 hours into the light period.
PolyA mRNA was purified from each sample every 2 hours for
the next 18 hours. The mRNA was pooled and used for cDNA
synthesis. The cDNA was directionally cloned into lambda
Zap II (Stratagene) in the EcoRI (5') and XhoI (3')
sites. pBluescript II SK- plasmids were excised from the
lambda ZAP clones by superinfection with Exassist
(Stratagene) phage. The library was normalized using
method 4 described in Bonaldo et al., (1996) Genome
Research 6: 791-806."
ORIGIN
Query Match 48.7%; Score 29.2; DB 13; Length 708;
Best local Similarity 69.0%; Pred. No. 8.1e+02;
Matches 40; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
QY 3 CATCAAGCAGGGCCCCCAAGAGCCCTTCGCGACTACGTGACCGCTTCTCAAGACC 60
| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 564 CTTCAACCGGAGACCCCATCGACCGCTTCATCTCTACTTCACCGCTACTCAAGGCC 621
RESULT 11
LOCUS CK165296 922 bp mRNA linear EST 05-DEC-2003
DEFINITION FGAS049245 Triticum aestivum FGAS: Talc7 Triticum aestivum cDNA,
mRNA sequence.
ACCESSION CK165296
VERSION CK165296
KEYWORDS CK165296.1 GI:38997204
SOURCE BST.
ORGANISM Triticum aestivum (bread wheat)
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticaceae; Triticum.
1 (bases 1 to 922)
```

[illegible]

the Human Genome Center, University of Tokyo Institute of Medical Science; plant material and RNA prepared at Texas A & M University; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.

Seq primer: Sugs (CTTCTGCTCTTAAGCTGCG).

Location/Qualifiers

1..781

/organism="Sorghum bicolor"

/mol\_type="mRNA"

/cultivar="Btx623"

/db\_xref="taxon:4558"

/clone="PH1\_26\_D03\_A002"

/lab\_host="DH10B-TI phage-resistant E. coli"

/note="Vector: pMB18S-FL3; Site 1: XhoI; Site 2: XhoI; The library was prepared from polyA+ RNA from Btx623 seedlings grown hydroponically in the absence of added phosphorous. At 14 days of age, roots and shoots were harvested and stored at -80 C until RNA was isolated. Double-stranded cDNA was cloned unidirectionally into different DraIII sites of the pMB18S-FL3 vector (5-prime DraIII site is CACTGTGTG, 3-prime DraIII site is CACCATGTG). XhoI excises the cDNA insert."

# ORIGIN

Query Match 48.3%; Score 29; DB 14; Length 781;  
Best Local Similarity 71.7%; Pred. No. 9.3e+02;  
Matches 38; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

OY 8 AGCAGGCCCCAAGAGAGCCCTCCGAGCTAGCTGAGCCGCTTCTTCAAGACC 60  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 580 AGCAGGACATCAAGAACCCCGCAGCGTCAACGAGACCCCATCTTCAAGAGC 632

## RESULT 13

AL811116 standard; mRNA; EST; 455 BP.

AC AL811116;

SV AL811116.1

DT 05-JUL-2002 (Rel. 72, Created)

DT 02-SEP-2002 (Rel. 72, last updated, Version 2)

DB Triticum aestivum (mercia) clone A10\_d37\_plate\_1 embryo (micropylar tip) 10

DB days post anthesis

XX EST; expressed sequence tag.

XX Triticum aestivum (bread wheat)

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;

OC Triticaceae; Triticum.

XX

RA Wilson I., Beswick R., Shepherd S., Barker G., Parker J., Owen P.,

RA Edwards D., Coghill J., Holdsworth M., Lenton J., Shewry P., Edwards K.;

RT "A BBSRC-funded wheat EST resource for the academic community";

RL Unpublished.

XX

RN

RP

RA Barker G.;

RL Submitted (03-JUL-2002) to the EMBL/GenBank/DBJ databases.

RL Barker G., Institute of Arable Crop Research, Long Ashton, Bristol BS41 9AF

RL United Kingdom.

Key Location/Qualifiers

PH source

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

1..455

/db\_xref="taxon:4565"

/mol\_type="mRNA"

/organism="Triticum aestivum"

/cultivar="mercia"

/dev\_stage="10 days post anthesis"

/tissue\_type="embryo (micropylar tip)"

/clone\_lib="d:37"

/clone="A10\_d37\_plate\_1"

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

1..455

/db\_xref="taxon:4565"

/mol\_type="mRNA"

/organism="Triticum aestivum"

/cultivar="mercia"

/dev\_stage="10 days post anthesis"

/tissue\_type="embryo (micropylar tip)"

/clone\_lib="d:37"

/clone="A10\_d37\_plate\_1"

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Sequence 45







```
Db 61 CTGGGCCCCGGGCAAGAGCACTATGCTGAAGCACTGTGTGGGCCAGCCGGAG 120
QY 121 CTGAGGGGCTTGGCCCTGGAACCCCGGCTGTGAGAGACCGCCGAGGGCTGCAAGCAGATC 180
Db 121 CTGAGGGGCTTGGCCCTGGAACCCCGGCTGTGAGAGACCGCCGAGGGCTGCAAGCAGATC 180
QY 181 ATGAGAGCAGCTGCAAGCCCGCTGCAAGACCGGCAACCGAGGAGCTGCGCAGCCTGTAAAC 240
Db 181 ATGAGAGCAGCTGCAAGCCCGCTGCAAGACCGGCAACCGAGGAGCTGCGCAGCCTGTAAAC 240
QY 241 ACCGTGGCCACCTGTACTGCTGCTGCAAGCCCGCATTCAGAGTCCGCCACACCAAGAGGCC 300
Db 241 ACCGTGGCCACCTGTACTGCTGCTGCAAGCCCGCATTCAGAGTCCGCCACACCAAGAGGCC 300
QY 301 CTGGAACAAGATCGAGAGAGAGCAAGAACAGTCCAGAGAGAACCCAGCAGGCCAAGAGAG 360
Db 301 CTGGAACAAGATCGAGAGAGAGCAAGAACAGTCCAGAGAGAACCCAGCAGGCCAAGAGAG 360
QY 361 GCCGACGGCAAGGTGAGCCAGAACCTACCTGTCAGAGAACCTGCAAGGGCCAGATGCTG 420
Db 361 GCCGACGGCAAGGTGAGCCAGAACCTACCTGTCAGAGAACCTGCAAGGGCCAGATGCTG 420
QY 421 CACCAAGCCATCAGCCCGCCGACCCCTGGAACGCTGTGGTGAAGGTGATGAGAGAGAGGCC 480
Db 421 CACCAAGCCATCAGCCCGCCGACCCCTGGAACGCTGTGGTGAAGGTGATGAGAGAGAGGCC 480
QY 481 TTCAAGCCCCGAGGTGATCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCAGAGAC 540
Db 481 TTCAAGCCCCGAGGTGATCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCAGAGAC 540
QY 541 CTGAACACGATGTGAACACCGTGGGGGCCACCAAGGCCCATGCAAGATGCTGAAGAGAC 600
Db 541 CTGAACACGATGTGAACACCGTGGGGGCCACCAAGGCCCATGCAAGATGCTGAAGAGAC 600
QY 601 ACCATCAACGAGAGAGCGCCGAGTGGACCGCCTGCAACCCCGTGCAAGCCCGCCCTG 660
Db 601 ACCATCAACGAGAGAGCGCCGAGTGGACCGCCTGCAACCCCGTGCAAGCCCGCCCTG 660
QY 661 GCCCCCGGCGAGTGGCGGACCCCGCGGAGGAGCATCGCCGCGGCAACAGCAACCTG 720
Db 661 GCCCCCGGCGAGTGGCGGACCCCGCGGAGGAGCATCGCCGCGGCAACAGCAACCTG 720
QY 721 CAGAGCAGATCGCTGTGATGACCAAGCAACCCCGCTGCGCGGCGACATCTTAAG 780
Db 721 CAGAGCAGATCGCTGTGATGACCAAGCAACCCCGCTGCGCGGCGACATCTTAAG 780
QY 781 CGGTGATCATCTGGGCTGGAACAAGATCGTGGAGATGAAGCCCGGTGAGCATCTG 840
Db 781 CGGTGATCATCTGGGCTGGAACAAGATCGTGGAGATGAAGCCCGGTGAGCATCTG 840
QY 841 GACATCCGCGAGGGCCCAAGAGAGCCCTTCGCGACTAGCTGAGCCGCTTCTTAAGACC 900
Db 841 GACATCCGCGAGGGCCCAAGAGAGCCCTTCGCGACTAGCTGAGCCGCTTCTTAAGACC 900
QY 901 CTGCGGCGGAGCAGGCCCAAGAGAGCCCTTCGCGACTAGCTGAGCCGCTTCTTAAGACC 960
Db 901 CTGCGGCGGAGCAGGCCCAAGAGAGCCCTTCGCGACTAGCTGAGCCGCTTCTTAAGACC 960
QY 961 CAGAACGCCAACCCCGACTGCAAGACCACTCTGCGCGCTCTGCGCCCGGCGCAACCTG 1020
Db 961 CAGAACGCCAACCCCGACTGCAAGACCACTCTGCGCGCTCTGCGCCCGGCGCAACCTG 1020
QY 1021 GAGGAGATGATACCGCTGCGGAGGGGCGGCGCGGCGCAAGGCGCGCGCTGCTG 1080
Db 1021 GAGGAGATGATACCGCTGCGGAGGGGCGGCGCGGCGCAAGGCGCGCGCTGCTG 1080
QY 1081 GCCGAGGCGATGAGCCAGGCCAACAGCGTGAACATCATGATGCAAGAGCAACTTCAAG 1140
Db 1081 GCCGAGGCGATGAGCCAGGCCAACAGCGTGAACATCATGATGCAAGAGCAACTTCAAG 1140
QY 1141 GGGCCCGGCGCAACGTCAAGTGTCTTCACTGCGGCAAGAGAGGCGCAACATCGCAAGAAC 1200
Db 1141 GGGCCCGGCGCAACGTCAAGTGTCTTCACTGCGGCAAGAGAGGCGCAACATCGCAAGAAC 1200
```

```
QY 1201 TGCCGCGCCCCCGCAAGAGAGGCTGTGTAAGTGGCGCAAGAGAGGCCACCAAGATGAAG 1260
Db 1201 TGCCGCGCCCCCGCAAGAGAGGCTGTGTAAGTGGCGCAAGAGAGGCCACCAAGATGAAG 1260
QY 1261 GACTGCAACCGAGCGCCAGGCCAACTTCTGTGGCAAGATCTGGCCCAAGAGGCGCC 1320
Db 1261 GACTGCAACCGAGCGCCAGGCCAACTTCTGTGGCAAGATCTGGCCCAAGAGGCGCC 1320
QY 1321 CCGGCACTTCTGTGAGAACCCGACAGAGCCCGCGCCCAACCGTGTCCCAACCGCC 1380
Db 1321 CCGGCACTTCTGTGAGAACCCGACAGAGCCCGCGCCCAACCGTGTCCCAACCGCC 1380
QY 1381 CCGGCGAGAGCTTCTGTGAGAACCCGACAGAGCCCGCGCCCAACCGTGTCCCAACAGAC 1440
Db 1381 CCGGCGAGAGCTTCTGTGAGAACCCGACAGAGCCCGCGCCCAACCGTGTCCCAACAGAC 1440
QY 1441 CGCAGCCCTACCGGAGCCCTGACCGCCCTGCGAGCCTGTGGGAGCGGCGCCCTG 1500
Db 1441 CGCAGCCCTACCGGAGCCCTGACCGCCCTGCGAGCCTGTGGGAGCGGCGCCCTG 1500
QY 1501 AGCCAGTAA 1509
Db 1501 AGCCAGTAA 1509
```

## RESULT 2

US-09-967-464-64

; Sequence 64, Application US/09967464

; Publication No. US20030138453A1

; GENERAL INFORMATION:

; APPLICANT: O'Hagan, Derek

; APPLICANT: Otten, Gillis

; APPLICANT: Donnelly, John J.

; APPLICANT: Polo, John M.

; APPLICANT: Barnett, Susan

; APPLICANT: Singh, Mamohan

; APPLICANT: Ulmer, Jeffrey

; APPLICANT: Dubensky, Jr., Thomas W.

; TITLE OF INVENTION: MICROPARTICLES FOR DELIVERY OF HETEROLOGOUS NUCLEIC ACIDS

; FILE REFERENCE: P16269.004

; CURRENT APPLICATION NUMBER: US/09/967,464

; PRIOR FILING DATE: 2002-04-11

; PRIOR APPLICATION NUMBER: 60/236,105

; PRIOR FILING DATE: 2000-09-28

; PRIOR APPLICATION NUMBER: 60/315,905

; PRIOR FILING DATE: 2001-08-30

; NUMBER OF SEQ ID NOS: 68

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 64

; LENGTH: 1509

; TYPE: DNA

; ORGANISM: Human immunodeficiency virus type 1

US-09-967-464-64

Query Match 99.9%; Score 1507.4; DB 10; Length 1509;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1508; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 ATGGGCGCGCGCGCGAGATCTGCGCGGAGAGAGTGAACAAGTGGAGAAATCCGC 60
Db 1 ATGGGCGCGCGCGCGAGATCTGCGCGGAGAGAGTGAACAAGTGGAGAAATCCGC 60
QY 61 CTGCGCGCGCGCGCGCAAGAAAGCACTACTAGCTGAAGCACTGTGTGGCGCAGCGGAG 120
Db 61 CTGCGCGCGCGCGCGCAAGAAAGCACTACTAGCTGAAGCACTGTGTGGCGCAGCGGAG 120
QY 121 CTGAGGGGCTTGGCCCTGGAACCCCGGCTGTGAGAGCCCGGAGGGCTGCAAGAGATC 180
Db 121 CTGAGGGGCTTGGCCCTGGAACCCCGGCTGTGAGAGCCCGGAGGGCTGCAAGAGATC 180
QY 181 ATGAAGCAGCTGCAAGCCCGCTGCAAGCCGCAACCGAGAGCTGCGAGCCTGTACAAC 240
Db 181 ATGAAGCAGCTGCAAGCCCGCTGCAAGCCGCAACCGAGAGCTGCGAGCCTGTACAAC 240
```

```

Db      181 ATGAAGCACTGACGCCGCCCTGCAGACCGGCAACCGAGAGCTGCCACCTGTACAC 240
QY      241 ACCGTGGCCACCCCTGTATCTGCTGCAACGCCGGCATCGAGTCCGCGACACCAAGAGGCC 300
Db      241 ACCGTGGCCACCCCTGTATCTGCTGCAACGCCGGCATCGAGTCCGCGACACCAAGAGGCC 300
QY      301 CTGACACAGATCGAGAGAGCAGAACAGATCCAGCAGAGAACCCAGAGGCCAAGAG 360
Db      301 CTGACACAGATCGAGAGAGCAGAACAGATCCAGCAGAGAACCCAGAGGCCAAGAG 360
QY      361 GCCGACGGCAAGTGAAGCCAGAACTACCCCATCTGTGAGAACTTGCAAGGCCAGATG 420
Db      361 GCCGACGGCAAGTGAAGCCAGAACTACCCCATCTGTGAGAACTTGCAAGGCCAGATG 420
QY      421 CACCAAGCCATCAGCCCCCGGCAACCTGAAGCCTTGGTGAAGGTGATCGAGAGAGGCC 480
Db      421 CACCAAGCCATCAGCCCCCGGCAACCTGAAGCCTTGGTGAAGGTGATCGAGAGAGGCC 480
QY      481 TTGAGCCCGGAGGTGATCCCATGTTTACCGCCCTGAGCGAGGGCGCCACCCCGAGAG 540
Db      481 TTGAGCCCGGAGGTGATCCCATGTTTACCGCCCTGAGCGAGGGCGCCACCCCGAGAG 540
QY      541 CTGAACACAGATGTTGAACACCGTGGGGCGGCCACAGGCCCGCATGCAATGCTGAAGAC 600
Db      541 CTGAACACAGATGTTGAACACCGTGGGGCGGCCACAGGCCCGCATGCAATGCTGAAGAC 600
QY      601 ACCATCAACGAGAGAGCGCCCGAGTGGACCGCTGACCCCGTGACAGGCCCGCCCTG 660
Db      601 ACCATCAACGAGAGAGCGCCCGAGTGGACCGCTGACCCCGTGACAGGCCCGCCCTG 660
QY      661 GCGCCCGGCGAGATGCGGACCCCGCGGCGAGCATCGCCGGCGCCACAGCAACCTG 720
Db      661 GCGCCCGGCGAGATGCGGACCCCGCGGCGAGCATCGCCGGCGCCACAGCAACCTG 720
QY      721 CAGGAGCAGATGCTGTGATGACCAAGAACCCCGGCGCGTGGCGCATCTACAAG 780
Db      721 CAGGAGCAGATGCTGTGATGACCAAGAACCCCGGCGCGTGGCGCATCTACAAG 780
QY      781 CGGTGATCATCTTGGGCTGAACAGATCTGCGGATGTAAGGCCCGTGAGCATCTG 840
Db      781 CGGTGATCATCTTGGGCTGAACAGATCTGCGGATGTAAGGCCCGTGAGCATCTG 840
QY      841 GACATCGCCAGAGGCCCGGCGAGGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGAC 900
Db      841 GACATCGCCAGAGGCCCGGCGAGGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGAC 900
QY      901 CTGCGCGCCGAGAGGCCCAACAGAGCGTGAAGAACTGATGACCGAGACCTGCTG 960
Db      901 CTGCGCGCCGAGAGGCCCAACAGAGCGTGAAGAACTGATGACCGAGACCTGCTG 960
QY      961 CAGAACGCAACCCCGACTGCAAGACCACTCTGCGCGCTCTCGGCCCGCGCCACCTG 1020
Db      961 CAGAACGCAACCCCGACTGCAAGACCACTCTGCGCGCTCTCGGCCCGCGCCACCTG 1020
QY      1021 GAGGAGATGATGACCGCTGCGCAGAGCGTGGGGCGGCCCGGCAAGGCCCGCGTCTG 1080
Db      1021 GAGGAGATGATGACCGCTGCGCAGAGCGTGGGGCGGCCCGGCAAGGCCCGCGTCTG 1080
QY      1081 GCGAGGCGATGAGCGCGCAACAGCGTGAACATCATGATGCAAGAGCAACTTCAAG 1140
Db      1081 GCGAGGCGATGAGCGCGCAACAGCGTGAACATCATGATGCAAGAGCAACTTCAAG 1140
QY      1141 GCGCCCGCGCGCAACGTCAAGTCTTCACTGCGCAAGAGGGCCCAATCGCAAGAAC 1200
Db      1141 GCGCCCGCGCGCAACGTCAAGTCTTCACTGCGCAAGAGGGCCCAATCGCAAGAAC 1200
QY      1201 TGCAGCGCCCCCGCAAGAGGGCTGCTGAAGTGCAGCAAGAGGGCCACAGATGAAG 1260
Db      1201 TGCAGCGCCCCCGCAAGAGGGCTGCTGAAGTGCAGCAAGAGGGCCACAGATGAAG 1260
QY      1261 GACTGCAACGAGCGCGCAAGCTTCTTGGGCAAGATCTGGCCAGCCCAAGAGGCCG 1320
Db      1261 GACTGCAACGAGCGCGCAAGCTTCTTGGGCAAGATCTGGCCAGCCCAAGAGGCCG 1320

```

```

QY      1321 CCGGCACTTCTGCAAGACCGCAGAGCCCGCGCCCGCCACCGTCCACCGCCCC 1380
Db      1321 CCGGCACTTCTGCAAGACCGCAGAGCCCGCGCCCGCCACCGTCCACCGCCCC 1380
QY      1381 CCGGCGAGAGCTTCCGCTTGCAGAGACCAACCCCGCCCGCCCAAGCAGAGCCCAAGAG 1440
Db      1381 CCGGCGAGAGCTTCCGCTTGCAGAGACCAACCCCGCCCGCCCAAGCAGAGCCCAAGAG 1440
QY      1441 CCGAGCCCTTACCGGAGCCCTGACCGCCCTGCGAGCCTGTTGGCAGCGGCCCTG 1500
Db      1441 CCGAGCCCTTACCGGAGCCCTGACCGCCCTGCGAGCCTGTTGGCAGCGGCCCTG 1500
QY      1501 AGCCAGTAA 1509
Db      1501 AGCCAGTAA 1509

```

```

RESULT 3
US-09-967-464-68
; Sequence 68, Application US/09967464
; Publication No. US20030138453A1
; GENERAL INFORMATION:
; APPLICANT: O'Hagan, Derek
; APPLICANT: Otten, Gillis
; APPLICANT: Donnelly, John J.
; APPLICANT: Polo, John M.
; APPLICANT: Barnett, Susan
; APPLICANT: Singh, Mamohan
; APPLICANT: Ulmer, Jeffrey
; APPLICANT: Dubenky, Jr., Thomas W.
; TITLE OR INVENTION: MICROPARTICLES FOR DELIVERY OF HETEROLOGOUS NUCLEIC ACIDS
; FILE REFERENCE: P16269.004
; CURRENT APPLICATION NUMBER: US/09/967,464
; PRIOR FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 60/236,105
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 60/315,905
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 68
; LENGTH: 1509
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
US-09-967-464-68

```

```

Query Match      99.7%; Score 1504.2; DB 10; Length 1509;
Best Local Similarity 99.2%; Pred. No. 2.2e-314;
Matches 1497; Conservative 12; Mismatches 0; Indels 0; Gaps 0;

QY      1 ATGGGCGCGCGCGCAGATCTTGGCGCGGAGAACTGACAAAGTGGAGAAATCCGC 60
Db      1 ATGGGCGCGCGCGCAGATCTTGGCGCGGAGAACTGACAAAGTGGAGAAATCCGC 60
QY      61 CTGCGCGCGCGCGCGCAGAGCACTACTATCTGAAGCACTGTGTGGGCCAGCGCGAG 120
Db      61 CTGCGCGCGCGCGCGCAGAGCACTACTATCTGAAGCACTGTGTGGGCCAGCGCGAG 120
QY      121 CTGAGAGGCTTGGCCCTGAACCCCGGCTGTGAGACCGCCGAGGGCTGCAAGCAGATC 180
Db      121 CTGAGAGGCTTGGCCCTGAACCCCGGCTGTGAGACCGCCGAGGGCTGCAAGCAGATC 180
QY      181 ATGAAGCAGCTGCAAGCCCGCTTGCAGACCGGCAACCGAGAGCTGCGCAGCTGTACAAC 240
Db      181 ATGAAGCAGCTGCAAGCCCGCTTGCAGACCGGCAACCGAGAGCTGCGCAGCTGTACAAC 240
QY      241 ACCGTGGCCACCCCTGTACTGCTGTCAGCGCCGCGCATCGAGTCCGCGACACCAAGAGGCC 300
Db      241 ACCGTGGCCACCCCTGTACTGCTGTCAGCGCCGCGCATCGAGTCCGCGACACCAAGAGGCC 300
QY      301 CTGACACAGATCGAGAGAGCAGAACAGATCCAGCAGAGAACCCAGAGGCCAAGAG 360

```



QY 481 TTCAAGCCCGAGGTGATCCCATGTTCAACCGCCCTGAGCGAGGGCCCAACCCCCAGAGAC 540  
DB 481 TTCAAGCCCGAGGTGATCCCATGTTCAACCGCCCTGAGCGAGGGCCCAACCCCCAGAGAC 540  
QY 541 CTGAACACGATGTTGAACACCGTGGGGCGGCGCAAGCGCCCATGAGATGCTGAAGGAC 600  
DB 541 CTGAACACGATGTTGAACACCGTGGGGCGGCGCAAGCGCCCATGAGATGCTGAAGGAC 600  
QY 601 ACCATCAACGAGGAGGCGCGCGAGTGGAGCCGCTGACCCCGTGCAGGCGCGCCCTG 660  
DB 601 ACCATCAACGAGGAGGCGCGCGAGTGGAGCCGCTGACCCCGTGCAGGCGCGCCCTG 660  
QY 661 GCGCCCGCGCGAGTGGCGGACCCCGCGCGAGCATGCGCGCGCGCAACGACCCCTG 720  
DB 661 GCGCCCGCGCGAGTGGCGGACCCCGCGCGAGCATGCGCGCGCGCAACGACCCCTG 720  
QY 721 CAGAGCAGATGCGCTGATGACCAAGCAACCCCGTGGCGCGGAGCATCTACAAG 780  
DB 721 CAGAGCAGATGCGCTGATGACCAAGCAACCCCGTGGCGCGGAGCATCTACAAG 780  
QY 781 CGGTGATCATCTGGGCTGAAACAAGATGTCGAGATGACAGCCCGTGAATCTG 840  
DB 781 CGGTGATCATCTGGGCTGAAACAAGATGTCGAGATGACAGCCCGTGAATCTG 840  
QY 841 GACATCCGCGAGGCGCGCGAGGCGCTTCCGCGAGTACGTCGACCGCTTCAAGAC 900  
DB 841 GACATCCGCGAGGCGCGCGAGGCGCTTCCGCGAGTACGTCGACCGCTTCAAGAC 900  
QY 901 CTGCGCGCGAGGCGCGCAACCGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 960  
DB 901 CTGCGCGCGAGGCGCGCAACCGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 960  
QY 961 CAGAACGCGCAACCGCGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1020  
DB 961 CAGAACGCGCAACCGCGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1020  
QY 1021 GAGAGATGATGACCGCTGCGAGGCGGCGGCGCGCGCGCGCGCGCGCGCGCG 1080  
DB 1021 GAGAGATGATGACCGCTGCGAGGCGGCGGCGCGCGCGCGCGCGCGCGCGCG 1080  
QY 1081 GCGAGGCGATGAGCGCGCAACCGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1140  
DB 1081 GCGAGGCGATGAGCGCGCAACCGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1140  
QY 1141 GCGCGCGCGCGCAACCGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1200  
DB 1141 GCGCGCGCGCGCAACCGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1200  
QY 1201 TGCGCGCGCGCGCAACCGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1260  
DB 1201 TGCGCGCGCGCGCAACCGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1260  
QY 1261 GACTGACGAGCGCGCAACCGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1320  
DB 1261 GACTGACGAGCGCGCAACCGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1320  
QY 1321 CCGCGCAACTTCTGCAAGACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380  
DB 1321 CCGCGCAACTTCTGCAAGACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380  
QY 1381 CCGCGCAAGCTTCCGCTTGAAGAGCAACCGCGCGCGCGCGCGCGCGCGCGCGCG 1440  
DB 1381 CCGCGCAAGCTTCCGCTTGAAGAGCAACCGCGCGCGCGCGCGCGCGCGCGCGCG 1440  
QY 1441 CCGGAGCGCTTACCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1500  
DB 1441 CCGGAGCGCTTACCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1500  
QY 1501 AGCCAGTAA 1509  
DB 1501 AGCCAGTAA 1509

RESULT 5  
US-09-899-575-51  
; Sequence 51, Application US/09899575  
; Publication No. US20030223961A1  
; GENERAL INFORMATION:  
; APPLICANT: Zur Megele, Jan  
; APPLICANT: Barnett, Susan W.  
; APPLICANT: Egnelbrecht, Susan  
; APPLICANT: van Rensburg, Estrelita Janse  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C  
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USRS THERBOF  
; FILE REFERENCE: PP01631.102  
; CURRENT APPLICATION NUMBER: US/09/899,575  
; CURRENT FILING DATE: 2001-07-05  
; PRIOR APPLICATION NUMBER: 09/475,704  
; PRIOR FILING DATE: 1999-12-30  
; NUMBER OF SEQ ID NOS: 135  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 51  
; LENGTH: 1494  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: HIV Type C  
US-09-899-575-51

Query Match 86.0%; Score 1297.6; DB 13; length 1494;  
Best Local Similarity 92.7%; Pred. No. 3.9e-270;  
Matches 1402; Conservative 0; Mismatches 89; Indels 21; Gaps 3;

QY 1 ATGGGCG 60  
DB 1 ATGGGCG 60  
QY 61 CTGCG 120  
DB 61 CTGCG 120  
QY 121 CTGAGGCGCTTCCGCTGAAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180  
DB 121 CTGAGGCGCTTCCGCTGAAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180  
QY 181 ATGAAGCAGTGCAGCG 240  
DB 181 ATGAAGCAGTGCAGCG 240  
QY 241 ACCGTGCGCACTCTGTACTGCTGTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300  
DB 241 ACCGTGCGCACTCTGTACTGCTGTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300  
QY 301 CTGCAACAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360  
DB 301 CTGCAACAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360  
QY 361 GCCGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420  
DB 361 GCCGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420  
QY 421 CACGAGCGCATCAGCG 480  
DB 421 CACGAGCGCATCAGCG 480  
QY 481 TTCAGCGCGGAGGTGATCCCATGTTCAACCGCCCTGAGCGAGGGCCCAACCCCAAGAC 540  
DB 481 TTCAGCGCGGAGGTGATCCCATGTTCAACCGCCCTGAGCGAGGGCCCAACCCCAAGAC 540  
QY 541 CTGAACACGATGTTGAACACCGTGGGGCGGCGCAAGCGCGCGCGCGCGCGCGCGCGCG 600  
DB 541 CTGAACACGATGTTGAACACCGTGGGGCGGCGCAAGCGCGCGCGCGCGCGCGCGCGCG 600  
QY 601 ACCATCAACGAGGAGGCGCGCGAGTGGAGCCGCTGACCCCGTGCAGGCGCGCGCGCG 660  
DB 601 ACCATCAACGAGGAGGCGCGCGAGTGGAGCCGCTGACCCCGTGCAGGCGCGCGCGCG 660



```

Db      601 ACCATCAACGAGAGAGCCCGCGAGTGGGACCGCACCCACCCCGTGCACCGCGCCCGTG 660
Qy      661 GCGCCCGGCGAGATGCGCGACCCCGCGGCGAGACATGCGCGGCCACAGCACCTTG 720
Db      661 GCGCCCGGCGAGATGCGCGAGCCCGCGGCGAGACATGCGCGGCCACAGCACCTTG 720
Qy      721 CAGGAGCAGATCGCTTGATGACCAACCCCGCGTGCCCGTGGCGACATCTACAAG 780
Db      721 CAGGAGCAGATCGCTTGATGACCAACCCCGCGTGCCCGTGGCGACATCTACAAG 780
Qy      781 CCGTGATCATCTGGGCTGAAACAGATGTCGCGATGTACAGCCCGTGAGCATCTG 840
Db      781 CCGTGATCATCTGGGCTGAAACAGATGTCGCGATGTACAGCCCGTGAGCATCTG 840
Qy      841 GACATCCGCGAGGCGCCCAAGAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 900
Db      841 GACATCAAGCAGGCGCCCAAGAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 900
Qy      901 CTGCGCGCGAGCAGGCGCCACCCAGAGACGTGAAGACTGATGACCGAGACCTTGCTGCTG 960
Db      901 CTGCGCGCGAGCAGGCGCCACCCAGAGACGTGAAGACTGATGACCGAGACCTTGCTGCTG 960
Qy      961 CAGAAAGCCAAACCCGACTGCAAGACATCTGCGCGCTCTGCGCCCGCGCGCACCTTG 1020
Db      961 CAGAAAGCCAAACCCGACTGCAAGACATCTGCGCGCTCTGCGCCCGCGCGCACCTTG 1020
Qy      1021 GAGAGATGATGACCGCTTGCGAGGCGTGCGCGCGCCCGCGCACAGGCGCGCTGCTG 1080
Db      1021 GAGAGATGATGACCGCTTGCGAGGCGTGCGCGCGCCCGCGCACAGGCGCGCTGCTG 1080
Qy      1081 GCGAGGCGATGACGAGGCGCAACGCTGACATCATGATGACAGAGCAACTTCAAG 1140
Db      1081 GCGAGGCGATGACGAGGCGCAACGCTGACATCATGATGACAGAGCAACTTCAAG 1140
Qy      1141 GCGCCCGCGCGCAAGTCAAGTCTTCACTGCGCGCAAGAGGCGCACTGCGCAAGAAC 1200
Db      1141 GCGAGCAACCGCATCATCAAGTCTTCACTGCGCGCAAGAGGCGCACTGCGCAAGAAC 1200
Qy      1201 TGCGCGCGCGCGCGCAAGAGGCTGCTGAAAGTGGCGCAAGAGGCGCACAGATGAAG 1260
Db      1201 TGCGCGCGCGCGCGCGCAAGAGGCTGCTGAAAGTGGCGCAAGAGGCGCACAGATGAAG 1260
Qy      1261 GACTGCAACCGAGCGCGCAAGGCGCACTTCTTGCGCAAGATCTGCGCGCACAGGCGCGC 1320
Db      1261 GACTGCAACCGAGCGCGCGCAAGGCGCACTTCTTGCGCAAGATCTGCGCGCACAGGCGCGC 1320
Qy      1321 CCGCGCAACTTCTTGCGCAAGACCGAGCGAGC---CGCGCGCGCGCACCGTGCGCGCAC 1377
Db      1321 CCGCGCAACTTCTTGCGCAAGACCGCGAGCGAGC---CGCGCGCGCGCACCGTGCGCGCAC 1377
Qy      1378 CCGCGCGCGAGGCTTCCGCTTGAGAGAGACACCCCGCGCGCGCGCGCGCGCGCGCGCG 1437
Db      1378 CCGCGCGCGAGGCTTCCGCTTGAGAGAGACACCCCGCGCGCGCGCGCGCGCGCGCGCG 1437
Qy      1438 GACCGGAGCGCTTACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1497
Db      1438 GACCGGAGCGCTTACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1497
Qy      1498 CTGAGCCAGTAA 1509
Db      1483 CTGAGCCAGTAA 1494

```

# RESULT 6 US-09-899-575-99

Sequence 99, Application US/09899575  
Publication No. US20030223961A1

## GENERAL INFORMATION:

APPLICANT: Zur Megede, Jan  
APPLICANT: Barnett, Susan W.  
APPLICANT: Egnelbrecht, Susan  
APPLICANT: van Rensburg, Estrelita Janse  
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C

```

; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP01631.102
; CURRENT APPLICATION NUMBER: US/09/899,575
; CURRENT FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 09/475,704
; PRIOR FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 99
; LENGTH: 1491
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Gag_TV2_C_ZAopt
US-09-899-575-99

```

Query Match 85.8%; Score 1294.8; DB 13; Length 1491;  
Best Local Similarity 92.7%; Pred. No. 1.6e-269;  
Matches 1401; Conservative 0; Mismatches 87; Indels 24; Gaps 3;

```

Qy      1 ATGGGCGCGCGCGCGAGCATCTGCGCGCGCGAGAACTGGACAAAGTGGAGAGATCCGC 60
Db      1 ATGGGCGCGCGCGCGAGCATCTGCGCGCGCGAGAACTGGACAAAGTGGAGATCCGC 60
Qy      61 CTGCGCGCGCGCGCGCAAGAGCACTACATGCTGAAGCACTTGTTGGGCCAGCCCGAG 120
Db      61 CTGCGCGCGCGCGCGCGCAAGCACTACATGCTGAAGCACTTGTTGGGCCAGCCCGAG 120
Qy      121 CTGAGAGGCTTCCGCTGAAACCCCGCGCTGCTGAGAACCGCGAGGCGCTGCAAGCATC 180
Db      121 CTGAGAGGCTTCCGCTGAAACCCCGCGCTGCTGAGAACCGCGAGGCGCTGCGCGCATC 180
Qy      181 ATGAAGCAGCTGAGCGCGCGCTGCAAGACCGCGAGAGAGCTGCGCGCATCTGTACAC 240
Db      181 ATCAAGCAGCTGAGCGCGCGCTGCAAGACCGCGAGAGAGCTGCGCGCATCTGTACAC 240
Qy      241 ACCGTGGCAACCTGTACTGCTGCGAGCGCGCGCATGAGGTCCGCGACACCAAGAGGCC 300
Db      241 ACCGTGGCAACCTGTACTGCTGCGAGCGCGCGCATGAGGTCCGCGACACCAAGAGGCC 300
Qy      301 CTGGAACAAGTCAAGAGAGAGCAACAAGTCCAGCAGAAAGCCAGAGGCCAGGCC 360
Db      301 CTGGAACAAGTCAAGAGAGAGCAACAAGTCCAGCAGAAAGCCAGAGGCCAGGCC 360
Qy      361 GCGCAGCGCAAGGTGAGCGCGAGAACTACCCCATCTGCAAGACCTGCAAGGCGCGAGATG 420
Db      361 GCGCAGCAAGAGGTGAGCGCGAGAACTACCCCATCTGCAAGACCTGCAAGGCGCGAGATG 420
Qy      421 CACGAGCGCATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
Db      421 CACGAGCGCATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
Qy      481 TTGAGCGCGAGGTATCCCATGTTCAACCGCGCTGAGCGAGGCGCGCACCCCGCAGGAC 540
Db      481 TTGAGCGCGAGGTATCCCATGTTCAACCGCGCTGAGCGAGGCGCGCACCCCGCAGGAC 540
Qy      541 CTGAACACGATGTTGAACAACGTTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
Db      541 CTGAACACGATGTTGAACAACGTTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
Qy      601 ACCATCAACGAGAGGCGCGCGAGTGGAGACCGCTGCAACCCCGTGACAGGCGCGCGCGTG 660
Db      601 ACCATCAACGAGAGGCGCGCGAGTGGAGACCGCTGCAACCCCGTGACAGGCGCGCGCGTG 660
Qy      661 GCGCCCGGCGAGATGCGCGAGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTG 720
Db      661 GCGCCCGGCGAGATGCGCGAGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTG 720
Qy      721 CAGGAGCAGATGCGCTGATGACCAAGCCCGCGTGCCCGTGGCGAGCATCTACAAG 780
Db      721 CAGGAGCAGATGCGCTGATGACCAAGCCCGCGTGCCCGTGGCGAGCATCTACAAG 780
Qy      781 CCGTGATCATCTGGGCTGAAACAGATGTCGCGATGTACAGCCCGTGAGCATCTCTG 840

```



```

Db      781 CGCTGATCATCTGGGCTGAAACAGATCGTGGCGCATGTACAGCCCCGTGACATCTG 840
QY      841 GACATCCGCGGAGGGCCCCAAGAGCCCTTCCCGACTACGTGACCGCTTCTTCAAGACC 900
Db      841 GACATCAAGCAGGGCCCCCAAGAGCCCTTCCCGACTACGTGACCGCTTCTTCAAGACC 900
QY      901 CTGCGCGCGGAGCAGGCGCAATCCCAAGAGCGTGAAGACTGTGATGACCGAGACCCCTGCTGTG 960
Db      901 CTGCGCGCGGAGCAGGCGCAATCCCAAGAGCGTGAAGACTGTGATGACCGAGACCCCTGCTGTG 960
QY      961 CAGAACCGCCAAACCCGACTGCAAGACCATCTGCGCGCTCTCGGCCCCCGCGCCACCTG 1020
Db      961 CAGAACCGCCAAACCCGACTGCAAGACCATCTGCGCGCTCTCGGCCCCCGCGCCACCTG 1020
QY      1021 GAGAGATGATGACCGCTGCAAGAGCGGCGTGGGCGGCGCCCGCAAGGCGCGCTGCTG 1080
Db      1021 GAGAGATGATGACCGCTGCAAGAGCGGCGTGGGCGGCGCCCGCAAGGCGCGCTGCTG 1080
QY      1081 GCGAGGCGCATGAGCCAGGCGCAACAGCGTGAATCATGATGCAAGAGCACTTCAAG 1140
Db      1081 GCGAGGCGCATGAGCCAGGCGCAACAGCGTGAATCATGATGCAAGAGCACTTCAAG 1140
QY      1141 GGGCCCCCGCGCAACGTCAAGTCTTCAACTGCGGCAAGAGGCGCACATCGCCAGAAG 1200
Db      1141 GGGCCCCCGCGCGCGCGTGAAGTCTTCAACTGCGGCGCGGCGGCGCACATCGCGCGAAC 1200
QY      1201 TGCGCGCGCGCGCGCGCAAGAGGCGCTGTGAAGTGCGGCAAGAGGCGCACAGATGAAG 1260
Db      1201 TGCGCGCGCGCGCGCGCAAGAGGCGCTGTGAAGTGCGGCAAGAGGCGCACAGATGAAG 1260
QY      1261 GACTGCAACGAGCGCGCAGGCGCACTTCTGCGGCAAGATCTGCGCCAGCCAGCAAGGCGCGC 1320
Db      1261 GACTGCAACGAGCGCGCAGGCGCACTTCTGCGGCAAGATCTGCGCCAGCCAGCAAGGCGCGC 1320
QY      1321 CCGCGCAACTTCTGCAAGACCGCAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1377
Db      1321 CCGCGCAACTTCTGCAAGACCGCAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1377
QY      1378 CCGCGCGCGCGGAGGCTTCCGCTTGAAGAGACCAACCCCGCGCGCGCGCGCGCGCGCGCGCG 1437
Db      1378 CCGCGCGCGCGGAGGCTTCCGCTTGAAGAGACCAACCCCGCGCGCGCGCGCGCGCGCGCGCG 1437
QY      1438 GACCGCGAGCCCTACCGCGAGCCCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1497
Db      1438 GACCGCGAGCCCTACCGCGAGCCCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1497
QY      1498 CTGAGCCAGTAA 1509
Db      1480 CTGAGCCAGTAA 1491

```

## RESULT 7

```

US-09-967-464-63
; Sequence 63, Application US/09967464
; Publication No. US20030138453A1
; GENERAL INFORMATION:
; APPLICANT: O'Hagan, Derek
; APPLICANT: Otten, Gillis
; APPLICANT: Donnelly, John J.
; APPLICANT: Polo, John M.
; APPLICANT: Barnett, Susan
; APPLICANT: Singh, Mamohan
; APPLICANT: Ulmer, Jeffrey
; APPLICANT: Dubensky, Jr., Thomas W.
; TITLE OF INVENTION: MICROPARTICLES FOR DELIVERY OF HETEROLOGOUS NUCLEIC ACIDS
; FILE REFERENCE: PP16269.004
; CURRENT APPLICATION NUMBER: US/09/967,464
; PRIOR FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 60/236,105
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 60/315,905
; PRIOR FILING DATE: 2001-08-30

```

```

; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 63
; LENGTH: 1479
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
US-09-967-464-63

Query Match      84.6%; Score 1276.8; DB 10; Length 1479;
Best Local Similarity 92.5%; Pred. No. 1.2e-265;
Matches 1399; Conservative 0; Mismatches 77; Indels 36; Gaps 4;

QY      1 ATGGCGCGCGCGCGCAGCATCTGCGCGCGGAGAGCTGGAGAGTGAGAGATCCGC 60
Db      1 ATGGCGCGCGCGCGCAGCATCTGCGCGCGGAGAGCTGGAGAGTGAGAGATCCGC 60
QY      61 CTGCGCGCGCGCGCGCAAGAGCACTACATGCTGAAGCACCTGTGTGGGCCAGCGCGAG 120
Db      61 CTGCGCGCGCGCGCGCAAGAGCTACATGATGAAGCACCTGTGTGGGCCAGCGCGAG 120
QY      121 CTGAGGCGCTTCCCTGAAACCCCGGCTGTGAGAGACCGCGAGGGCTGCAAGCATC 180
Db      121 CTGAGAGGTTCCCTGAAACCCCGGCTGTGAGAGACCGCGAGGGCTGCAAGCATC 180
QY      181 ATGAAGCAGCTGACGCCCGCTGACAGACCGGCAACGAGAGCTGCGCAGCTGTACAAC 240
Db      181 ATCCGCAAGTCAACCCCGCTGACAGACCGGCAACGAGAGCTGCGCAGCTGTACAAC 240
QY      241 ACCGTGCGCAACCTGTACTGCTGTGACGCGCGCATGAGGTCCGCGACCAAGAGGCC 300
Db      241 ACCGTGCGCAACCTGTACTGCTGTGACGCGCGCATGAGGTCCGCGACCAAGAGGCC 300
QY      301 CTGACAAGATCGAGAGAGAGAGCAAGATGCCAGAGAGACCCAGCAGGCGCAAGAG 360
Db      301 CTGACAAGATCGAGAGAGAGAGCAAGATGCCAGAGAGATCCAGCAGGCGCAAGAG 360
QY      361 GCGGAC--GCAAGGTGAGCCAGAACTACCCCATGCTGTGAGAACTGACGCGCAGATG 417
Db      361 GCGGAC--GCAAGGTGAGCCAGAACTACCCCATGCTGTGAGAACTGACGCGCAGATG 417
QY      418 GTGCAACGAGCCATCAGCCCGCGCAACCTGAACGCTGTGTGAAGTGAAGAGAG 477
Db      418 GTGCAACGAGCCATCAGCCCGCGCAACCTGAACGCTGTGTGAAGTGAAGAGAGAG 477
QY      478 GCCTTACGCGCGGAGTATCCCATATGTTCAACCGCCCTGAGCGAGGGCGCGCGCGCGCG 537
Db      478 GCCTTACGCGCGGAGTATCCCATATGTTCAACCGCCCTGAGCGAGGGCGCGCGCGCGCG 537
QY      538 GACCTGAACAGATGTTGAACACCGTGGCGGCGCACAGCGCGCATGCAATGCTGAAG 597
Db      538 GACCTGAACAGATGTTGAACACCGTGGCGGCGCACAGCGCGCATGCAATGCTGAAG 597
QY      541 GACCTGAACAGATGTTGAACACCGTGGCGGCGCACAGCGCGCATGCAATGCTGAAG 600
Db      541 GACCTGAACAGATGTTGAACACCGTGGCGGCGCACAGCGCGCATGCAATGCTGAAG 600
QY      598 GACACCATCAAGAGAGGCGCGGAGTGGAGCGGCTGACCCCGTGACGCGCGCGCGCG 657
Db      598 GACACCATCAAGAGAGGCGCGGAGTGGAGCGGCTGACCCCGTGACGCGCGCGCGCG 657
QY      601 GACACCATCAAGAGAGGCGCGGAGTGGAGCGGCTGACCCCGTGACGCGCGCGCGCG 660
Db      601 GACACCATCAAGAGAGGCGCGGAGTGGAGCGGCTGACCCCGTGACGCGCGCGCGCG 660
QY      658 GTGCGCGCGCGCGCAGATGCGCGAGCCCGCGCGCAGCGACATCGCGCGCGCGCACAG 717
Db      658 GTGCGCGCGCGCGCAGATGCGCGAGCCCGCGCGCAGCGACATCGCGCGCGCGCACAG 717
QY      661 ATGCGCGCGCGCGCAGATGCGCGAGCGCGCGCGCAGCGACATCGCGCGCGCGCACAG 720
Db      661 ATGCGCGCGCGCGCAGATGCGCGAGCGCGCGCGCAGCGACATCGCGCGCGCGCACAG 720
QY      718 CTGCAAGAGAGATGCGCTGTGATGACCAAGCAACCCCGTGCGCGCGCGCATCTAC 777
Db      718 CTGCAAGAGAGATGCGCTGTGATGACCAAGCAACCCCGTGCGCGCGCGCATCTAC 777
QY      721 CTGCAAGAGAGATGCGCTGTGATGACCAAGCAACCCCGTGCGCGCGCGCATCTAC 780
Db      721 CTGCAAGAGAGATGCGCTGTGATGACCAAGCAACCCCGTGCGCGCGCGCATCTAC 780
QY      778 AAGCGGTGATCATCTGCGCTGAAACAAGATGTGCGGATGTACAGCCCGGTGACATC 837
Db      778 AAGCGGTGATCATCTGCGCTGAAACAAGATGTGCGGATGTACAGCCCGGTGACATC 837
QY      838 CTGACATCCGCGCAGGCGCGCGCAAGAGCGCTTCCGCGACTACGTTGAGCCGTTCTCAAG 897
Db      838 CTGACATCCGCGCAGGCGCGCGCAAGAGCGCTTCCGCGACTACGTTGAGCCGTTCTCAAG 897
QY      841 CTGACATCAAGCAGGCGCGCGCAAGAGCGCTTCCGCGACTACGTTGAGCCGTTCTCAAG 900
Db      841 CTGACATCAAGCAGGCGCGCGCAAGAGCGCTTCCGCGACTACGTTGAGCCGTTCTCAAG 900
QY      898 ACCCTGCGCGCGCGAGCGCAAGCGCAACCGAGAGCTGAAGACTGATGACCGAGACCTGCTG 957

```

||||| 901 ACCCTGCGCGCGAGAGAGACCCAGAGAGTGAAGAACTGATGACCGACACCTGCTG 960  
QY 958 GTGCAAGACGCAACCCGACTGCAAGACCATCTGCGCGCTCTGCGCCCGCGCCACC 1017  
DB 961 GTGCAAGACGCAACCCGACTGCAAGACCATCTGCGCGCTCTGCGCCCGCGCCACC 1020  
QY 1018 CTGAGAGAGATGATGACCGCTGCGAGAGCGCTGCGCGCGCCCGCCCAAGGCCCGCTG 1077  
DB 1021 CTGAGAGAGATGATGACCGCTGCGAGAGCGCTGCGCGCGCCCGCCCAAGGCCCGCTG 1080  
QY 1078 CTGCGCGAGCGATGAGCCAGCGCAAGCGTGAACATCATGATGAGAGAGCACTTC 1137  
DB 1081 CTGCGCGAGCGATGAGCCAGCGCAAGCGTGAACATCATGATGAGAGAGCACTTC 1137  
QY 1138 AAGGCGCCCGCGCGCAAGCTCAAGTCTTCAACTGCGCGCAAGAGAGGCCCATCGCCAG 1197  
DB 1138 AAGGCGCCCGCGCGCATCTCAAGTCTTCAACTGCGCGCAAGAGAGGCCCATCGCCAG 1197  
QY 1198 AACTGCGCGCGCGCGCGCAAGAGGCTGCTGAAAGTGGCGCAAGAGAGGCCCATGATG 1257  
DB 1198 AACTGCGCGCGCGCGCGCATCTCAAGTCTTCAACTGCGCGCAAGAGAGGCCCATGATG 1257  
QY 1258 AAGGACTGCAACCGAGCGCGCAAGCTTCTGCGCAAGATCTGCGCGCAAGAGAGGCC 1317  
DB 1258 AAGGACTGCAACCGAGCGCGCAAGCTTCTGCGCAAGATCTGCGCGCAAGAGAGGCC 1317  
QY 1318 CGCGCGCGCAACTTCTGCAAGACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1377  
DB 1318 CGCGCGCGCAACTTCTGCAAGACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1359  
QY 1378 CCCCCCGCGAGAGCTTCTGAGAGAGACACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1437  
DB 1360 CCCCCCGCGAGAGCTTCTGAGAGAGACACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1419  
QY 1438 GACCGCGAGCGCTTCTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1497  
DB 1420 G-----ACCGAGAGACCTGACCAAGCTGAAAGAGCTGTTGCGCAAGCGCGCG 1467  
QY 1498 CTGAGCCAGTAA 1509  
DB 1468 CTGAGCCAGTAA 1479

## RESULT 8

US-09-899-575-3

; Sequence 3, Application US/09899575

; Publication No. US2003023961A1

; GENERAL INFORMATION:

; APPLICANT: Zur Megele, Jan

; APPLICANT: Barnett, Susan W.

; APPLICANT: Egnelbrecht, Susan

; APPLICANT: van Rensburg, Estrelita Janse

; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C

; FILE REFERENCE: PP01631.102

; CURRENT APPLICATION NUMBER: US/09/899,575

; CURRENT FILING DATE: 2001-07-05

; PRIOR APPLICATION NUMBER: 09/475,704

; PRIOR FILING DATE: 1999-12-30

; NUMBER OF SEQ ID NOS: 135

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 3

; LENGTH: 1479

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: synthetic Gag

; US-09-899-575-3

Query Match 84.6%; Score 1276.8; DB 13; Length 1479;  
Best Local Similarity 92.5%; Pred. No. 1.2e-265;

Matches 1399; Conservative 0; Mismatches 77; Indels 36; Gaps 4;  
QY 1 ATGGCGCGCGCGCGCGAGCATCTGCGCGCGCGCGAGAGCTGCAAGTGGAGAGATCCGC 60  
DB 1 ATGGCGCGCGCGCGCGAGCATCTGCGCGCGCGCGAGAGCTGCAAGTGGAGAGATCCGC 60  
QY 61 CTGCGCGCGCGCGCGAGAGCACTATGCTGAGACCTGTGTGGCGCGAGCCCGAG 120  
DB 61 CTGCGCGCGCGCGCGAGAGAGTGTCTATGATGAGACCTGTGTGGCGCGAGCCCGAG 120  
QY 121 CTGAGAGGCTTGGCTTGAACCCCGCGCTGCTGAGACCGCGAGGCTGCAAGCATTC 180  
DB 121 CTGAGAGAGTTCGCTTGAACCCCGCGCTGCTGAGACCGAGGCTGCAAGCATTC 180  
QY 181 ATGAGCAGCTGAGCGCGCGCTGAGACCGAGAGAGTGGAGCTGCAAGCTGTAAAC 240  
DB 181 ATCGGCGAGCTGAGCGCGCGCTGAGACCGAGAGAGTGGAGCTGCAAGCTGTAAAC 240  
QY 241 ACCGTGGCGCGCGCTGATCTGCTGAGCGCGCGCGAGTCCCGGACCAAGAGGCC 300  
DB 241 ACCGTGGCGCGCGCTGATCTGCTGAGCGCGCGCGAGTCCCGGACCAAGAGGCC 300  
QY 301 CTGCAAGATGAGAGAGAGAGAGCAAGTCCAGAGAGAGAGAGAGAGAGAGAGAGAG 360  
DB 301 CTGCAAGATGAGAGAGAGAGAGCAAGTCCAGAGAGAGAGAGAGAGAGAGAGAGAG 360  
QY 361 GCGGAC---GGCAAGTGAAGCGAGAACTACCCCATCTGCAAGACTGCAAGGCGCGATG 417  
DB 361 GCGGACAAAGGAGAGAGTGAAGCGAGAACTACCCCATCTGCAAGACTGCAAGGCGCGATG 420  
QY 418 GTGACCAAGGCGCATCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 477  
DB 421 GTGACCAAGGCGCATCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480  
QY 478 GCTTCAAGCGCGCGAGTATCCCATGTTCAACCGCGCTGAGCGAGGCGCGCGCGCGCG 537  
DB 481 GCTTCAAGCGCGCGAGTATCCCATGTTCAACCGCGCTGAGCGAGGCGCGCGCGCGCG 540  
QY 538 GACTGAACAGATGTTGAACACCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 597  
DB 541 GACTGAACAGATGTTGAACACCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600  
QY 598 GACACCATCAACGAGAGAGCGCGCGAGTGGAGCGCGCTGCAACCGCGTGAAGCGCGCG 657  
DB 601 GACACCATCAACGAGAGAGCGCGCGAGTGGAGCGCGCTGCAACCGCGTGAAGCGCGCG 660  
QY 658 GTGCGCGCGCGCGCGAGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 717  
DB 661 ATGCGCGCGCGCGCGAGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720  
QY 718 CTGCAAGAGCAGATGCTGCTGATGACCAAGACCGCGCGCGCGCGCGCGCGCGCGCG 777  
DB 721 CTGCAAGAGCAGATGCTGCTGATGACCAAGACCGCGCGCGCGCGCGCGCGCGCGCG 780  
QY 778 AAGCGGTGATCTCTGCGCTGGAACAAGATCGTGGAGATGTACAGCGCGCGCGCGCATC 837  
DB 781 AAGCGGTGATCTCTGCGCTGGAACAAGATCGTGGAGATGTACAGCGCGCGCGCGCATC 840  
QY 838 CTGCAATCCCG 897  
DB 841 CTGCAATCAAGCG 900  
QY 898 ACCGTGGCGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 957  
DB 901 ACCGTGGCGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 960  
QY 958 GTGCAAGACGCAACCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1017  
DB 961 GTGCAAGACGCAACCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020  
QY 1018 CTGAGAGAGATGATGACCGCTGCGAGAGCGCGCTGCGCGCGCGCGCGCGCGCGCG 1077  
DB 1021 CTGAGAGAGATGATGACCGCTGCGAGAGCGCGCTGCGCGCGCGCGCGCGCGCGCG 1080

QY	1078	CTGGCCGAGGCGATGAGCCAGGCGCAACAGCTGAACATCATGATGCAGAAAGCAACTTC	1137
Db	1081	CTGGCCGAGGCGATGAGCCAGGCGCAACACC---AGCGTGAATGATGCAGAAAGCAACTTC	1137
QY	1138	AAGGCCCCCGGCGCAACGTCAAGTGTCTTCACTGCGGCAAGAGGGCCAATCGCCAAAG	1197
Db	1138	AAGGCCCCCGGCGCATGTGTCAAGTGTCTTCACTGCGGCAAGAGGGCCAATCGCCCGC	1197
QY	1198	AACCTGCGCGCCCCCGGCAAGAGGGCTGTCTGAAGTGCAGAGGCGCACAGATG	1257
Db	1198	AACCTGCGCGCCCCCGGCAAGAGGGCTGTCTGAAGTGCAGAGGCGCACAGATG	1257
QY	1258	AAGGACTGCACCGAGCGCCAGGCCCAACTTCTGTGGGCAAGATCTTGCCCAAGAGGC	1317
Db	1258	AAGGACTGCACCGAGCGCCAGGCCCAACTTCTGTGGGCAAGATCTTGCCCAAGAGGC	1317
QY	1318	CGCCCCGGGCAACTTCTGTGCAAGCCGACGAGCCCGCGCCCCACCGTGCCACCGCC	1377
Db	1318	CGCCCCGGGCAACTTCTGTGCAAGCGG-----CGCCGAGCCACCGCC	1359
QY	1378	CCCCCGGCGGAGGCTTCCGCTTGAGGAGACCACCCCCGCCCCAAGGAGAGGCCAAG	1437
Db	1360	CCCCCGGCGGAGGCTTCCGCTTGAGGAGACCACCCCCGGCCAAGGAGGAGAGGCAAG	1419
QY	1438	GACCGCGAGCCCTAACCGCGAGCCCTGACCGCCCTGCGGACCTGTTCCGCAAGCGGCCCC	1497
Db	1420	G-----ACCGCGAGACCTGACCGACCTGAAAGAGCTGTTCCGCAACGACCCC	1467
QY	1498	CTGAGCCAGTTAA 1509	
Db	1468	CTGAGCCAGTTAA 1479	

```

RESULT 9
US-10-190-435-18
; Sequence 18, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; PIR REFERENCE: P18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 3162
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: Description of Artificial Sequence: GagRtmnt_C
US-10-190-435-18

```

Query Match	84.5%	Score 1275.2;	DB 15;	length 3162;
Best Local Similarity	92.5%;	Pred. No. 2,6e-265;		
Matches 1398;	Conservative	0;	Mismatches 78;	Indels 36;
			Gaps	4;
Qy	1	ATGGGCGCCCGCGCCAGCATCCCTGCGCGCGCGGAGAGCTGAGCAAGTGGAGAGATCCGC	60	
Db	7	ATGGGCGCCCGCGCCAGCATCTGCGCGCGCGGAGAGCTGAGCCTCTGGAGCGCATCCGC	66	
Qy	61	CTGCGCCCGCGCGGCAAGAGCACTACATGCTGAAGCACTGTGTGTGGCCAGCCGGAG	120	
Db	67	CTGCGCCCGCGCGGCAAGAGTGTCTACATGATGAGCACTGTGTGTGGCCAGCCGGAG	126	
Qy	121	CTGAGAGGCTTCGCCCTGAACCCCGCCCTGCTTGAGAGCCGCCAGGGCTGCAAGCAGATC	180	

Db	127	CTGGAGAAAGTTCGCCCTTGAAACCCCGGCTCTGTCGAGACCAGCGAGGGCTGCAAGCAGATC	186
Oy	181	ATGAAGCAGCTGACAGCCCGCCTTGACAGCCGGCAACCGAGAGCTGCGCAGCTGTACAAAC	240
Db	187	ATCCGCCAGCTGCACCCCGCCTTGACAGACCGGCACGGAGAGCTGAMGAGCCTGTTC AAC	246
Oy	241	ACCGTGGCCACCCTGTACTGCGTGCACGCGCGCATGAGGTCGCGCACCAAGGAGGCC	300
Db	247	ACCGTGGCCACCCTGTACTGCGTGCACGAGAAAGATGAGGTCGCGCACCAAGGAGGCC	306
Oy	301	CTGACAAAGATCGAGGAGGACAGAACAAAGTCCAGCAGAAAGATCCAGCAGGCCAAAGAG	360
Db	307	CTGACAAAGATCGAGGAGGACAGAACAAAGTCCAGCAGAAAGATCCAGCAGGCCAAAGAG	366
Oy	361	GCCGAC--GGCAAGGTGAGCCAGAACTAACCCCATGCTGTGCAGAACTTGCAAGGCCAGATG	417
Db	367	GCCGACAAAGGCGAAGGTGAGCCAGAACTAACCCCATGCTGTGCAGAACTTGCAAGGCCAGATG	426
Oy	418	GTGCACACGAGCCATCAGCCCCCGCACCCCTTGACCGCCTTGAGGTGAAGGTGATCGAGGAGAG	477
Db	427	GTGCACACGAGCCATCAGCCCCCGCACCCCTTGACCGCCTTGAGGTGAAGGTGATCGAGGAGAG	486
Oy	478	GCCTTCAAGCCCGAGGTGATCCCATGTTTCAACCGCCCTGAGCGAGGGGCGCCACCCCCAG	537
Db	487	GCCTTCAAGCCCGAGGTGATCCCATGTTTCAACCGCCCTGAGCGAGGGGCGCCACCCCCAG	546
Oy	538	GACCTGAACAAGATGTTGAACAACGTGGGCGGGCCACCAAGCGGCCATGACAGATGCTGAAG	597
Db	547	GACCTGAACAAGATGTTGAACAACGTGGGCGGGCCACCAAGCGGCCCATGACAGATGCTGAAG	606
Oy	598	GACACCATCAACGAGGAGCGCGCGAGTGGGACCGCTTGCAACCCCGTGCAAGGCCCGGCC	657
Db	607	GACACCATCAACGAGGAGCGCGCGAGTGGGACCGCTTGCAACCCCGTGCAAGGCCCGGCC	666
Oy	658	GTGGCCCCCGGCCAGATGCGCGCAACCCCGCGCAGCGACATGCGCGCGGCCACCAAGACCC	717
Db	667	ATGCCCCCGGCCAGATGCGCGCAACCCCGCGCAGCGACATGCGCGCGGCCACCAAGACCC	726
Oy	718	CTGCAGAGAGAGATGCGCTGAGATGACCAAGCAACCCCGGTGCCCCGTGGGCGCAGATCTAC	777
Db	727	CTGCAGAGAGAGATGCGCTGAGATGACCAAGCAACCCCGCATCTCCCGTGGGCGCAGATCTAC	786
Oy	778	AAGCGGTGATCATCTTGGGCTTGAAACAAGATCGTGCGGATGTACAGCCCGGTGAGCATC	837
Db	787	AAGCGGTGATCATCTTGGGCTTGAAACAAGATCGTGCGGATGTACAGCCCGGTGAGCATC	846
Oy	838	CTGACATCCGCCAAGGGCCCCCAAGAGCCCTTCCGCGACTACGTGGAACCGCTTCTCAAG	897
Db	847	CTGACATCAAGCAAGGGCCCCCAAGAGCCCTTCCGCGACTACGTGGAACCGCTTCTCAAG	906
Oy	898	ACCCTGCGCGCGAGCAGGCCAACCCAGAGCGTGAAGAACTGGAATGACCGAGACCTGCTG	957
Db	907	ACCCTGCGCGCGAGCAGGCCAACCCAGAGCGTGAAGAACTGGAATGACCGAGACCTGCTG	966
Oy	958	GTGCAGAACGCGCAACCCCGACTGCAGAACCATCTGCGCGCTCTCGGCCCGCGCCACACC	1017
Db	967	GTGCAGAACGCGCAACCCCGACTGCAGAACCATCTGCGCGCTCTCGGCCCGCGCCACACC	1026
Oy	1018	CTGAGAGGAGATGATGACCGCTGCGCAGGGCGTGGGGCGGCCCCCGGCCACAAGGCCCGCTG	1077
Db	1027	CTGAGAGGAGATGATGACCGCTGCGCAGGGCGTGGGGCGGCCCCCGGCCACAAGGCCCGCTG	1086
Oy	1078	CTGGCCGAGCGGATGAGCCAGGGCCACAGCGGTGAACATCATGATGCAGAGAGCAACTTC	1137
Db	1087	CTGGCCGAGCGGATGAGCCAGGGCCACAGCGGTGAACATCATGATGCAGAGAGCAACTTC	1143
Oy	1138	AAGGGCCCCCGCGCAACGTCAAGTGTCTTCAACTGCGGCAAGAGAGGGCCACATCGCCAAAG	1197
Db	1144	AAGGGCCCCCGCGCATGTCTCAAGTGTCTTCAACTGCGGCAAGAGAGGGCCACATCGCCCGC	1203
Oy	1198	AACTGCCGCGCCCCCGCAAGAGGGCTGCTGGAAGTGCAGGCAAGGAGGCCACCAAGATG	1257
Db	1204	AACTGCCGCGCCCCCGCAAGAGGGCTGCTGGAAGTGCAGGCAAGGAGGCCACCAAGATG	1263













Query Match 84.5%; Score 1275.2; DB 15; Length 4606;  
 Best Local Similarity 92.5%; Pred. No. 2.5e-265;  
 Matches 1398; Conservative 0; Mismatches 78; Indels 36; Gaps 4;

```

OY 1 ATGGGCGCCCGCCAGCATCTTCGCGCGGCGAGAACTGACAACTGGAGAAATCCGC 60
DB 3128 ATGGGCGCCCGCCAGCATCTTCGCGCGGCGAGAACTGACAACTGGAGAAATCCGC 3187
OY 61 CTGCGCCCGCGCGCAAGAGCACTACATGCTGAAGCACTGTGTGGCCAGCCGCGAG 120
DB 3188 CTGCGCCCGCGCGCAAGAGCTACATGATGAAGCACTGTGTGGCCAGCCGCGAG 3247
OY 121 CTGAGAGGCTTCGCTTGAACCCCGGCTGTGAGAACCGCGAGGGCTGCAAGCATC 180
DB 3248 CTGAGAGAGTTCGCTTGAACCCCGGCTGTGAGAACCGCGAGGGCTGCAAGCATC 3307
OY 181 ATGAAGCAGCTGACGCCCGCTGACAGCCGCGACCGAGAGCTGCGAGCTGTACAAC 240
DB 3308 ATCGGCAAGCTGACGCCCGCTGACAGCCGCGACCGAGAGCTGAAGAGCTGTACAAC 3367
OY 241 ACCGTGGCCACCTGTACTGCGGTGACGCCCGCATTCGAGTCCGCGACCAAGAGGCC 300
DB 3368 ACCGTGGCCACCTGTACTGCGGTGACGCCCGCATTCGAGTCCGCGACCAAGAGGCC 3427
OY 301 CTGGAACAAGATCGAGAGAGAGCAAAAGTCCGACAGAGACCGACAGGCCAAGAG 360
DB 3428 CTGGAACAAGATCGAGAGAGAGCAAAAGTCCGACAGAGATCGACAGGCCAAGAG 3487
OY 361 GCGGAC--GGAAGGTGAGCCAGAACTACCCCATCTGTGAGAACTGTGAGGGCCAGATG 417
DB 3488 GCGGACAAAGGCAAGGTGAGCCAGAACTACCCCATCTGTGAGAACTGTGAGGGCCAGATG 3547
OY 418 GTGCAACAGGCAATCAGCCCGCGCACCTGAAAGCTGTGAGTGAAGTGAAGAGAG 477
DB 3548 GTGCAACAGGCAATCAGCCCGCGCACCTGAAAGCTGTGAGTGAAGTGAAGAGAG 3607
OY 478 GCTTCAGCCCGCGAGGTGATCCCATGTTCAACGCTGTGAGCGAGCGCCACCCCGAG 537
DB 3608 GCTTCAGCCCGCGAGGTGATCCCATGTTCAACGCTGTGAGCGAGCGCCACCCCGAG 3667
OY 538 GACCTGAACAAGATGTTGAACAACGTTGGCGCGCAACAGGCCGCAATGACAGATCTGAAG 597
DB 3668 GACCTGAACAAGATGTTGAACAACGTTGGCGCGCAACAGGCCGCAATGACAGATCTGAAG 3727
OY 598 GACACCATCAACGAGAGGCGCGGAGTGGAGCGGCTGACCCCGTGCAGGCGCGGCC 657
DB 3728 GACACCATCAACGAGAGGCGCGGAGTGGAGCGGCTGACCCCGTGCAGGCGGCC 3787
OY 658 GTGGCCCCCGCGCAGATGCGCGAACCCCGCGCGCAGCATGCGCGCGCCACGACGACC 717
DB 3788 ATGCGCCCCCGCGCAGATGCGCGAACCCCGCGCGCAGCATGCGCGCGCCACGACGACC 3847
OY 718 CTGCAAGAGAGATGCGCTGATGACGAAACCCCGCTGCGCTGGCGGACATCTAC 777
DB 3848 CTGCAAGAGAGATGCGCTGATGACGAAACCCCGCTGCGCTGGCGGACATCTAC 3907
OY 778 AAGCGGTGATCATCTGGGCTGAAACAAGATCGTGGGATGTACAGCCCGCTGACATC 837
DB 3908 AAGCGGTGATCATCTGGGCTGAAACAAGATCGTGGGATGTACAGCCCGCTGACATC 3967
OY 838 CTGGAATCCCGCGAGGGGCGCGAGAGCCCTTCCCGGACTACGTGACCGCTTCTCAAG 897
DB 3968 CTGGAATCCCGCGAGGGGCGCGAGAGCCCTTCCCGGACTACGTGACCGCTTCTCAAG 4027
OY 898 ACCCTGCGCGCGCGAGAGGCGCAACCGAGAGCTGAAGAACTGATGACCGAGACCTCTG 957
DB 4028 ACCCTGCGCGCGCGAGAGGCGCAACCGAGAGCTGAAGAACTGATGACCGAGACCTCTG 4087
OY 958 GTGCAAGAGCGCAACCCCGAGCTGCAAGAACCTCTGGCGGCTCTGCGGCCCGCGCGCAC 1017
DB 4088 GTGCAAGAGCGCAACCCCGAGCTGCAAGAACCTCTGGCGGCTCTGCGGCCCGCGCGCAC 4147
OY 1018 CTGGAAGAGATGATGACCGCTGCGAGGGCGGTGGCGGCCCGCGCAAGGCCCGCGTG 1077

```

```

DB 4148 CTGAGAGAGATGATGACCGCTGCGAGGGCGGTGGCGGCCCGCCAGCCACAAGGCCCGGTG 4207
OY 1078 CTGCGCCGAGCGCATGAGCCAGGCCCAACAGCGTGAACATCATGTAGAGAGCAACTTC 1137
DB 4208 CTGCGCCGAGCGCATGAGCCAGGCCCAACAC--ACGTGATGATGAGAGAGCAACTTC 4264
OY 1138 AAGGCCCCCGCGCAAGTGAAGTCTTCAACTGCGGCAAGAGGGCCACATCGCCAAAG 1197
DB 4265 AAGGCCCCCGCGCATGTGAAGTCTTCAACTGCGGCAAGAGGGCCACATCGCCCGC 4324
OY 1198 AACTGCGCGCGCGCGCAAGAGGGCTGTGGAAGTGCAGAGAGGGCCACAGATG 1257
DB 4325 AACTGCGCGCGCGCGCAAGAGGGCTGTGGAAGTGCAGAGAGGGCCACAGATG 4384
OY 1258 AAGGACTGACCGAGCGCGCAAGCTTCTGCGCAAGATCTGCGCCAGCCACAAGGCC 1317
DB 4385 AAGGACTGACCGAGCGCGCAAGCTTCTGCGCAAGATCTGCGCCAGCCACAAGGCC 4444
OY 1318 CGCCCCCGCACTTCTGAGAACCGGACGAGGCCCGCGCCCAAGCTGCGCACCGCC 1377
DB 4445 CGCCCCCGCACTTCTGAGAGCG-----CGCCGAGCCACCGCC 4486
OY 1378 CCCCCCGCGAGACTTCCGCTTGAAGAGACCAACCCCGCCCGCCAGAGAGGCCAAG 1437
DB 4487 CCCCCCGCGAGACTTCCGCTTGAAGAGACCAACCCCGCGCGAGAGAGAGCAAG 4546
OY 1438 GACCGCGAGCCCTACCGCGAGCCCTGACCGCCCTGCGGAGCTGTTCGCGAGCGCCCC 1497
DB 4547 G-----ACCGGAGACCTGACCAAGCTTGAAGAGCTGTTCGCGCAACGACCCC 4594
OY 1498 CTGAGCCCAATTA 1509
DB 4595 CTGAGCCCAATTA 4606

```

RESULT 15  
 US-10-190-435-36  
 ; Sequence 36, Application US/10190435  
 ; Publication No. US20030143248A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ZUR MEGRDE, Jan  
 ; APPLICANT: BARNETT, Susan W.  
 ; APPLICANT: LIAN, Ying  
 ; APPLICANT: ENGELBRECHT, Susan  
 ; APPLICANT: VAN RENSBURG, Estrelita J.  
 ; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C  
 ; FILE REFERENCE: P18133.003 / 2302-18133  
 ; CURRENT APPLICATION NUMBER: US/10/190,435  
 ; CURRENT FILING DATE: 2002-12-30  
 ; NUMBER OF SEQ ID NOS: 319  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 36  
 ; LENGTH: 4615  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURES:  
 ; OTHER INFORMATION: Description of Artificial Sequence:  
 ; OTHER INFORMATION: gp160mod.TV1.dv2-gagmod.BW965  
 US-10-190-435-36

Query Match 84.5%; Score 1275.2; DB 15; Length 4615;  
 Best Local Similarity 92.5%; Pred. No. 2.5e-265;  
 Matches 1398; Conservative 0; Mismatches 78; Indels 36; Gaps 4;

```

OY 1 ATGGGCGCCCGCGCAGCATCTTCGCGCGGCGAGAACTGACAACTGGAGAAATCCGC 60
DB 3137 ATGGGCGCCCGCGCAGCATCTTCGCGCGGCGAGAACTGACAACTGGAGAAATCCGC 3196
OY 61 CTGCGCCCGCGCGCAAGAGCACTACATGCTGAAGCACTGTGTGGCCAGCGCGAG 120
DB 3197 CTGCGCCCGCGCGCAAGAGTGTACTATGATGAAGCACTGTGTGGCCAGCGCGAG 3256

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

## OM nucleic - nucleic search, using SW model

Run on: June 1, 2004, 10:26:25 ; Search time 612 Seconds  
(without alignments)  
10474.724 Million cell updates/sec

Title: US-09-475-704A-4  
Perfect score: 1509  
Sequence: 1 atggcgccgcgcgcgcagcat.....ggcgccctgagccagtaa 1509

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_29Jan04:\*

- 1: geneseqn1980s:\*
- 2: geneseqn1990s:\*
- 3: geneseqn2000s:\*
- 4: geneseqn2001as:\*
- 5: geneseqn2001bs:\*
- 6: geneseqn2002as:\*
- 7: geneseqn2003as:\*
- 8: geneseqn2003bs:\*
- 9: geneseqn2003cs:\*
- 10: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1509	100.0	1509	3	AAA51610
2	1509	100.0	1509	6	AA144549 HIV synch
3	1509	100.0	1509	6	ABL39954
4	1491.4	98.8	1509	6	ABL39958
5	1489.8	98.7	1509	3	AAA51626
6	1489.8	98.7	1509	6	ABL44553
7	1297.6	85.0	1494	6	ABL39972
8	1294.8	85.8	1491	6	ABL40020
9	1276.8	84.6	1479	3	AAA51609
10	1276.8	84.6	1479	6	AA144548
11	1276.8	84.6	1479	6	ABL39953
12	1275.2	84.5	3162	9	ADCl3239
13	1275.2	84.5	3462	9	ADCl3237
14	1275.2	84.5	4419	9	ACA03523
15	1275.2	84.5	4419	9	ADCl3240
16	1275.2	84.5	4615	9	ADCl3257
17	1275.2	84.5	4702	9	ADCl3259
18	1275.2	84.5	4716	7	ACA03522
19	1275.2	84.5	4716	9	ADCl3238
20	1273.6	84.4	2742	7	ACA03524
21	1273.6	84.4	2742	9	ADCl3241
22	1273.6	84.4	3930	9	ADCl3230
23	1273.6	84.4	3930	9	ADCl3231

24	1273.6	84.4	3930	9	ADCl3232	ADCl3232 DNA of HI
25	1273.6	84.4	5145	7	ACA03521	ACA03521 Synthetic
26	1273.6	84.4	5145	9	ADCl3233	ADCl3233 DNA of HI
27	1272.6	84.3	4713	7	ACA03592	ACA03592 Synthetic
28	1272.6	84.3	4713	9	ADCl3280	ADCl3280 DNA of HI
29	1271.6	84.3	5184	7	ACA03591	ACA03591 Synthetic
30	1271.6	84.3	5184	9	ADCl3279	ADCl3279 DNA of HI
31	1270.2	84.2	2742	7	ACA03590	ACA03590 Synthetic
32	1270.2	84.2	2742	9	ADCl3278	ADCl3278 DNA of HI
33	1260.8	83.6	1479	3	AAA51625	AA144552 HIV codon
34	1259.2	83.4	1479	6	ABL39957	ABL39957 Synthetic
35	1234.4	81.8	4546	9	ADCl3255	AA144552 HIV-1 p55.
36	1234.4	81.8	4546	9	ADCl3255	ADCl3255 DNA of HI
37	1202.2	79.7	4288	4	AAH20868	AAH20868 HIV-1 sub
38	1201	79.6	1515	3	AAA70412	AAA70412 Synthetic
39	1201	79.6	4472	3	AAA70472	AAa70472 HIV bicis
40	1201	79.6	4608	3	AAA70473	AAa70473 HIV bicis
41	1201	79.6	4689	3	AAA70471	AAa70471 HIV bicis
42	1201	79.6	4766	3	AAA70470	AAa70470 HIV bicis
43	1199.4	79.5	2799	7	ACC78493	ACC78493 HIV GagTa
44	1199.4	79.5	2799	7	ACC78493	ACC78493 HIV GagTa
45	1199.4	79.5	3205	7	ACC78492	ACC78492 HIV GagRT

## ALIGNMENTS

RESULT 1  
AAA51610  
ID AAA51610 standard; DNA; 1509 BP.  
XX  
AC AAA51610;  
XX  
DT 31-OCT-2000 (first entry)  
XX  
DE HIV synthetic Gag polynucleotide.  
XX  
KW Gag; expression cassette; antigenic; type C; HIV; Env; synthetic;  
KW DNA immunization; packaging cell line; antigen presentation; ss.  
XX  
OS Human immunodeficiency virus; type C strain AF110967.  
OS Synthetic.  
XX  
PN WO200039304-A2.  
XX  
PD 06-JUL-2000.  
XX  
PF 30-DEC-1999; 99WO-US031273.  
XX  
PR 31-DEC-1998; 98US-0114495P.  
PR 01-SEP-1999; 99US-0152195P.  
PA (CHIR ) CHIRON CORP.  
PI Barnett S, Zur Megede J;  
XX WPI; 2000-452401/39.  
XX  
DR  
XX Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV Env  
PT polypeptide and the polypeptide useful for immunizing a mammal especially  
PT human against HIV.  
XX  
XX  
PS Claim 2; Page 93; 113pp; English.  
XX  
XX Expression cassettes comprising a polynucleotide encoding antigenic type  
CC C human immunodeficiency virus (HIV) Gag or Env polypeptides are useful  
CC in DNA immunization, generation of packaging cell lines and production of  
CC Gag- and/or Env-containing proteins. Synthetic Env and Gag expression  
CC cassettes exhibit increased potency for induction of cytotoxic T-  
CC lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1 self-  
CC assemble into non-infectious virus-like particles which are used as a  
CC matrix for the proper presentation of an antigen entrapped or associated  
CC to the immune system of the host



XX Sequence 1509 BP; 320 A; 556 C; 472 G; 161 T; 0 U; 0 Other;

Query Match 100.0%; Score 1509; DB 3; Length 1509;

Best Local Similarity 100.0%; Pred. No. 3.8e-177;

Matches 1509; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGGGCGCGCGCCAGCATCTCGCGCGGAGCTGACCAAGTGGAGAGATCCGC 60
DB 1 ATGGGCGCGCGCCAGCATCTCGCGCGGAGCTGACCAAGTGGAGAGATCCGC 60
QY 61 CTGCGCGCGCGCGCGAGAGCACTACATGCTGAAGCACTGTGTGTGGCCAGCCGAG 120
DB 61 CTGCGCGCGCGCGCGAGAGCACTACATGCTGAAGCACTGTGTGTGGCCAGCCGAG 120
QY 121 CTGAGAGGCTTCCGCTGAACCCCGGCTGTGTGAAGACCGCGGAGGCTGCAAGCAGATC 180
DB 121 CTGAGAGGCTTCCGCTGAACCCCGGCTGTGTGAAGACCGCGGAGGCTGCAAGCAGATC 180
QY 181 ATGAAGCAGCTGCAAGCCCGCTGCAAGCCCGCAACGAGAGCTGCGCAGCCTGTACAAC 240
DB 181 ATGAAGCAGCTGCAAGCCCGCTGCAAGCCCGCAACGAGAGCTGCGCAGCCTGTACAAC 240
QY 241 ACCGTGGCCACCTGTACTGTCTGCAAGCCCGCATGAGTCCGCGACCAAGAGGCC 300
DB 241 ACCGTGGCCACCTGTACTGTCTGCAAGCCCGCATGAGTCCGCGACCAAGAGGCC 300
QY 301 CTGGAACAAGTGAAGAGAGAGAGCAAGTCCAGAGAGAGAGAGAGAGAGAGAGAG 360
DB 301 CTGGAACAAGTGAAGAGAGAGAGAGCAAGTCCAGAGAGAGAGAGAGAGAGAGAGAG 360
QY 361 GCCGACGGCAAGTGAAGCCAGAACTACCCCATCTGTGAGAACTGCAAGGCGCAGATG 420
DB 361 GCCGACGGCAAGTGAAGCCAGAACTACCCCATCTGTGAGAACTGCAAGGCGCAGATG 420
QY 421 CACCAAGGCAATCAGCCCGCGCACTGAAACGCTGTGAGAGTGAAGAGAGAGAG 480
DB 421 CACCAAGGCAATCAGCCCGCGCACTGAAACGCTGTGAGAGTGAAGAGAGAGAG 480
QY 481 TTCAAGCCCGAGGTGATCCCATGTTCAACCGCCTGAGCGAGGCGCCACCCCGCAGAC 540
DB 481 TTCAAGCCCGAGGTGATCCCATGTTCAACCGCCTGAGCGAGGCGCCACCCCGCAGAC 540
QY 541 CTGAACACAGATGTTGAACACCGTGGGCGGCAACAGGCGCGCATGCAAGTGTGAAGAC 600
DB 541 CTGAACACAGATGTTGAACACCGTGGGCGGCAACAGGCGCGCATGCAAGTGTGAAGAC 600
QY 601 ACCATCAACAGAGAGGCGCGCGAGTGGAGCGGCTGCAACCCGCTGCAAGGCGCGCG 660
DB 601 ACCATCAACAGAGAGGCGCGCGAGTGGAGCGGCTGCAACCCGCTGCAAGGCGCGCG 660
QY 661 GCCCGCGCGCGAGATGCGGCAACCCCGCGGAGGAGCATGCGCGCGCAACGACCCCTG 720
DB 661 GCCCGCGCGCGAGATGCGGCAACCCCGCGGAGGAGCATGCGCGCGCAACGACCCCTG 720
QY 721 CAGGAGCAGATGCGCTGTGATGACCAAGACCCCGCTGTGGCGGAGCATCTACAAG 780
DB 721 CAGGAGCAGATGCGCTGTGATGACCAAGACCCCGCTGTGGCGGAGCATCTACAAG 780
QY 781 CGGTGATCATCTGTGGGCTGAAACAAGATGTGGGATGTACAGCCCGGTGAGCATCTCTG 840
DB 781 CGGTGATCATCTGTGGGCTGAAACAAGATGTGGGATGTACAGCCCGGTGAGCATCTCTG 840
QY 841 GACATCCGCGAGAGGCGCGCAAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 900
DB 841 GACATCCGCGAGAGGCGCGCAAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 900
QY 901 CTGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
DB 901 CTGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
QY 961 CAGAACGCGAACCAGCTGCAAGACCATCTGTGGCGCTGTGGCGCGCGCGCAGCCCTG 1020
DB 961 CAGAACGCGAACCAGCTGCAAGACCATCTGTGGCGCTGTGGCGCGCGCGCAGCCCTG 1020

```

```

DB 961 CAGAACGCGAACCAGCTGCAAGACCATCTGTGGCGCTGTGGCGCGCGCGCAGCCCTG 1020
QY 1021 GAGGAGATGATGACCGCTTCCAGGGCGGTGGCGGCGCGCGCAAGAGCGCGTGTCTG 1080
DB 1021 GAGGAGATGATGACCGCTTCCAGGGCGGTGGCGGCGCGCGCGCAAGAGCGCGTGTCTG 1080
QY 1081 GCCGAGGCGATGAGCCAGGCGCAACAGCGTGAACATCATGATTCAGAGAGCACTTCAAG 1140
DB 1081 GCCGAGGCGATGAGCCAGGCGCAACAGCGTGAACATCATGATTCAGAGAGCACTTCAAG 1140
QY 1141 GCGCGCGCGCGCAACGTCAAGTCTTCAACTGCGCGCAAGAGGCGCACATCGCGCAAGAAC 1200
DB 1141 GCGCGCGCGCGCAACGTCAAGTCTTCAACTGCGCGCAAGAGGCGCACATCGCGCAAGAAC 1200
QY 1201 TGCCGCGCGCGCGCAAGAGAGGCTGTGAAAGTCCGCGCAAGAGGCGCACATGATGAAG 1260
DB 1201 TGCCGCGCGCGCGCAAGAGAGGCTGTGAAAGTCCGCGCAAGAGGCGCACATGATGAAG 1260
QY 1261 GACTGCAACCGAGCGCCAGGCGCAACTTCTGTGGCAAGATCTGCGCCAGCAAGGCGCGC 1320
DB 1261 GACTGCAACCGAGCGCCAGGCGCAACTTCTGTGGCAAGATCTGCGCCAGCAAGGCGCGC 1320
QY 1321 CCGCGCAACTTCTGTGCAAGACCGCAAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
DB 1321 CCGCGCAACTTCTGTGCAAGACCGCAAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
QY 1381 CCGCGCGAGAGCTTCCGCTTGAAGAGAGACCAACCCCGCGCGCGCGCGCGCGCGCGCG 1440
DB 1381 CCGCGCGAGAGCTTCCGCTTGAAGAGAGACCAACCCCGCGCGCGCGCGCGCGCGCGCG 1440
QY 1441 CGCGAGCCCTTACCGCGAGAGCGCGCTGACCGCGCTGCGAGCGCTGTGCGAGCGCGCG 1500
DB 1441 CGCGAGCCCTTACCGCGAGAGCGCGCTGACCGCGCTGCGAGCGCTGTGCGAGCGCGCG 1500
QY 1501 AGCCAGTAA 1509
DB 1501 AGCCAGTAA 1509

```

## RESULT 2

AAL44549

ID AAL44549 standard; DNA; 1509 BP.

XX

AAL44549;

XX

29-AUG-2003 (revised)

DT 08-NOV-2002 (first entry)

XX

HIV-1 p55gag polypeptide coding sequence 2.

DE

XX

HIV, ds; vaccine; gene; immune response; microparticle;

adsorbent surface; poly(alpha-hydroxy acid); polyhydroxy butyric acid;

polycaprolactone; polyorthoester; polycyanacrylate; detergent;

submicron emulsion; viral infection; bacterial infection;

parasitic infection; HIV-1 p55gag polypeptide.

XX

Human immunodeficiency virus 1.

OS

XX

WO200226209-A2.

XX

04-APR-2002.

XX

28-SEP-2001; 2001WO-US030540.

XX

28-SEP-2000; 2000US-0236105P.

XX

30-AUG-2001; 2001US-0315905P.

XX

(CHIR ) CHIRON CORP.

XX

O'hagan D, Otten G, Donnelly JJ, Polo JM, Barnett S, Singh M;

XX

Ulmer J, Dubensky TW;

XX

WPI; 2002-519084/55.

XX

XX A microparticle to which a biologically active macromolecule is adsorbed,  
PT for use as a vaccine composition to treat viral, bacterial or parasitic  
PT infections, comprises a polymer microparticle, a detergent and a  
PT submicron emulsion.  
XX  
XX Claim 72; Fig 2; 100pp; English.  
XX  
XX The invention relates to a method of raising an immune response in a host  
CC animal. The method of the invention comprises administering a  
CC microparticle that has an adsorbent surface to which a first biologically  
CC active macromolecule (e.g. a nucleic acid) has been adsorbed. The  
CC microparticle comprises a polymer microparticle of poly(alpha-hydroxy  
CC acid), a polyhydroxy butyric acid, a polycaprolactone, a polyorthoester,  
CC a polycyanacrylate, a detergent, and submicron emulsion. The method/  
CC microparticle of the invention is useful for immunising a host animal  
CC against viral, bacterial or parasitic infections. The present DNA  
CC sequence encodes a HIV-1 p55gag polypeptide. (Updated on 29-AUG-2003 to  
CC standardise OS field)  
XX  
SQ Sequence 1509 BP; 320 A; 556 C; 472 G; 161 T; 0 U; 0 Other;  
Query Match 100.0%; Score 1509; DB 6; Length 1509;  
Best Local Similarity 100.0%; Pred. No. 3.8e-177;  
Matches 1509; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGGCGCCCGCGCGAGCATCTCGCGCGGAGAACTGGACAAGTGGAGAAATCCGC 60  
DB 1 ATGGGCGCCCGCGCGAGCATCTCGCGCGGAGAACTGGACAAGTGGAGAAATCCGC 60  
QY 61 CTGCGCCCGCGCGCGAGAGCACTACATCTGAAGCACCTGTGTGGCCAGCCGCGAG 120  
DB 61 CTGCGCCCGCGCGCGAGAGCACTACATCTGAAGCACCTGTGTGGCCAGCCGCGAG 120  
QY 121 QNAGAGGCTTCGCCCTGAACCCCGCTGCTGAGAACCCCGAGGCTGCAAGCAGATC 180  
DB 121 CTGAGGCTTCGCCCTGAACCCCGCTGCTGAGAACCCCGAGGCTGCAAGCAGATC 180  
QY 181 ATGAAGCAGCTGCAAGCCCGCTGCAAGCCGCGACCCGAGAGCTGCGAGCTGTACAC 240  
DB 181 ATGAAGCAGCTGCAAGCCCGCTGCAAGCCGCGACCCGAGAGCTGCGAGCTGTACAC 240  
QY 241 ACCGTGGCCACCTGTACTGCTGTCAGCCGCGCATCGAGTCCGCGACCAAGAGGCC 300  
DB 241 ACCGTGGCCACCTGTACTGCTGTCAGCCGCGCATCGAGTCCGCGACCAAGAGGCC 300  
QY 301 CTGCAACAAGATCGAGAGAGCAGAAACAAGTCCAGCAGAGAACCAGAGGCCAAGAG 360  
DB 301 CTGCAACAAGATCGAGAGAGCAGAAACAAGTCCAGCAGAGAACCAGAGGCCAAGAG 360  
QY 361 GCCGAGCGCAAGGTGAGCGCAACTACCCCATCTGTCAGAACTGCAAGGCCAGATG 420  
DB 361 GCCGAGCGCAAGGTGAGCGCAACTACCCCATCTGTCAGAACTGCAAGGCCAGATG 420  
QY 421 CACCAAGCCATCAAGCCCGCGCAACCTGAACGCTGAGTGAAGTGAAGAGGCC 480  
DB 421 CACCAAGCCATCAAGCCCGCGCAACCTGAACGCTGAGTGAAGTGAAGAGGCC 480  
QY 481 TTCAAGCCCGAGGTATCCCATATGTTCAACGCGCTGAGGAGGGCGCAACCCCAAG 540  
DB 481 TTCAAGCCCGAGGTATCCCATATGTTCAACGCGCTGAGGAGGGCGCAACCCCAAG 540  
QY 541 CTGAACAGATGTTGAACAACGCTGGCGCGCAACAGGCGCATGACATGCTGAAGGAC 600  
DB 541 CTGAACAGATGTTGAACAACGCTGGCGCGCAACAGGCGCATGACATGCTGAAGGAC 600  
QY 601 ACCATCAACGAGAGCGCGCGAGTGGACCGGCTGCAACCCCGTGCAGGCGCGCCG 660  
DB 601 ACCATCAACGAGAGCGCGCGAGTGGACCGGCTGCAACCCCGTGCAGGCGCGCCG 660  
QY 661 GCCCGCGCGCAGATGCGCGACCCCGCGCGCAGCAGATGCGCGCGCAGCAACCTG 720  
DB 661 GCCCGCGCGCAGATGCGCGACCCCGCGCGCAGCAGATGCGCGCGCAGCAACCTG 720

QY 721 CAGAGCAGATCGCCTGTGATGACCAAGCAACCCCCCGTGCCTGGGCGACATCTACAAG 780  
DB 721 CAGAGCAGATCGCCTGTGATGACCAAGCAACCCCCCGTGCCTGGGCGACATCTACAAG 780  
QY 781 CGGTGATCATCTCTGGGCTGAAACAAGATGTGCGGATGTACAGCCCCGTGAGCATCTC 840  
DB 781 CGGTGATCATCTCTGGGCTGAAACAAGATGTGCGGATGTACAGCCCCGTGAGCATCTC 840  
QY 841 GACATCCGCGAGGGCCCCCAAGAGCCCTTCCGACTACGTTGACCGCTTCTTCAAGACC 900  
DB 841 GACATCCGCGAGGGCCCCCAAGAGCCCTTCCGACTACGTTGACCGCTTCTTCAAGACC 900  
QY 901 CTGCGCGCGAGCAGGCGCAACCAAGACGTGAAGAACTGATGACCGAGACCTGCTGTG 960  
DB 901 CTGCGCGCGAGCAGGCGCAACCAAGACGTGAAGAACTGATGACCGAGACCTGCTGTG 960  
QY 961 CAGAACGCCAACCCCGACTGCAAGACCATCTGCGGCTCTGCGCCCCGCGCAACCTG 1020  
DB 961 CAGAACGCCAACCCCGACTGCAAGACCATCTGCGGCTCTGCGCCCCGCGCAACCTG 1020  
QY 1021 GAGAGATGATGACCGCTGCGCAAGGCGGTGGCGCCCCGCGCAAGGCGCGTGTCTG 1080  
DB 1021 GAGAGATGATGACCGCTGCGCAAGGCGGTGGCGCCCCGCGCAAGGCGCGTGTCTG 1080  
QY 1081 GCCGAGCGCATGAGCCAGGCGCAACAGCTGAACATCATGATGCAAGAGCAACTTCAAG 1140  
DB 1081 GCCGAGCGCATGAGCCAGGCGCAACAGCTGAACATCATGATGCAAGAGCAACTTCAAG 1140  
QY 1141 GCGCCCGCGCGCAAGCTCAAGTGTCTTCACTGCGGCAAGAGGCGCAATGCGCAAGAAC 1200  
DB 1141 GCGCCCGCGCGCAAGCTCAAGTGTCTTCACTGCGGCAAGAGGCGCAATGCGCAAGAAC 1200  
QY 1201 TGCGCGCGCCCCCGCAAGAGGCTGTGAAGTGGCGGCAAGAGGCGCAACGATGAAG 1260  
DB 1201 TGCGCGCGCCCCCGCAAGAGGCTGTGAAGTGGCGGCAAGAGGCGCAACGATGAAG 1260  
QY 1261 GACTGCAACGAGCGCCAGGCGCAACTTCTGCGCAAGATCTGCGCCAGCAAGGGCGCG 1320  
DB 1261 GACTGCAACGAGCGCCAGGCGCAACTTCTGCGCAAGATCTGCGCCAGCAAGGGCGCG 1320  
QY 1321 CCGCGCAACTTCTGCAAGACCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380  
DB 1321 CCGCGCAACTTCTGCAAGACCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380  
QY 1381 CCGCGGAGGCTTCCGCTTGAAGAGACCAACCCCGCGCGCGCGCGCGCGCGCGCG 1440  
DB 1381 CCGCGGAGGCTTCCGCTTGAAGAGACCAACCCCGCGCGCGCGCGCGCGCGCGCG 1440  
QY 1441 CGCGAGCCCTACCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1500  
DB 1441 CGCGAGCCCTACCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1500  
QY 1501 AGCCAAGTAA 1509  
DB 1501 AGCCAAGTAA 1509  
RESULT 3  
ABL39954  
ID ABL39954 standard; DNA; 1509 BP.  
XX  
XX ABL39954;  
XX AC  
XX 15-MAY-2002 (first entry)  
XX DT  
XX  
XX Synthetic Gag polynucleotide sequence SEQ ID NO:4.  
XX DB  
XX Human immunodeficiency virus type C; antigenic HIV type C protein;  
XX immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;  
XX immunostimulant; gene therapy; gene; ds.  
XX KW  
XX Human immunodeficiency virus; type C.  
OS

OS	Synthetic.
XX	
PN	WO200204493-A2.
XX	
PD	17-JAN-2002.
XX	
PP	05-JUL-2001; 2001WO-US021241.
XX	
PR	05-JUL-2000; 2000US-00610313.
XX	
PA	(CHIR ) CHIRON CORP.
PA	(UYST-) UNIV STELLENBOSCH.
XX	
PI	Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg EJ;
XX	
DR	WPI; 2002-154920/20.
XX	
PT	New polynucleotides encoding antigenic HIV Type C polypeptides, useful in
PT	applications including DNA immunization or generation of packaging cell
PT	lines, particularly in gene therapy.

Example 1; Fig 2; 233pp; English.

The present invention describes expression cassettes comprising a polynucleotide sequence encoding a polypeptide comprising immunogenic HIV type C polypeptides. The expression cassettes comprise any of the HIV type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or Nef (I). (I) have immunostimulant activity and can be used in gene therapy. The HIV type C polynucleotides are useful in applications including DNA immunisation, generation of packaging cell lines, and production of HIV type C proteins. The polynucleotides are particularly useful in gene therapy and DNA immunisation applications. ABI39942 to ABI40054 and ABB06204 to ABB06215 represent sequences used in the exemplification of the present invention

Sequence 1509 BP; 320 A; 556 C; 472 G; 161 T; 0 U; 0 Other;

Query Match	100.0%;	Score 1509;	DB 6;	Length 1509;
Best Local Similarity	100.0%;	Pred. No. 3.8e-177;		
Matches 1509; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

QY	1	ATGCGCGCGCCGCGCAGCATCTCTGCGCGCGGAGAAAGCTTGACCAAGTGGAGAAAGATCCGC	60
Db	1	ATGCGCGCGCCGCGCAGCATCTCTGCGCGCGGAGAAAGCTTGACCAAGTGGAGAAAGATCCGC	60
QY	61	CTGCGCGCGCGCGGCAAGAAAGCACTAATGCTGAAGCACTGTGTGTGGCCAGCCCGAG	120
Db	61	CTGCGCGCGCGCGGCAAGAAAGCACTAATGCTGAAGCACTGTGTGTGGCCAGCCCGAG	120
QY	121	CTGAGAGGCTTCGCGCTTGAAACCCGCGCTGTGGAAGACCGCGAGGCTGCAGCAGATC	180
Db	121	CTGAGAGGCTTCGCGCTTGAAACCCGCGCTGTGGAAGACCGCGAGGCTGCAGCAGATC	180
QY	181	ATGAAGCAGCTGCAGCCCGCGCTGCAGACCGGACCGAGAGCTGCGCAGCTGTACAAC	240
Db	181	ATGAAGCAGCTGCAGCCCGCGCTGCAGACCGGACCGAGAGCTGCGCAGCTGTACAAC	240
QY	241	ACCGTGGCCACCTGTACTGCGTGCAACCGCGCATCGAGGTCCGCGACCAAGAGGCGC	300
Db	241	ACCGTGGCCACCTGTACTGCGTGCAACCGCGCATCGAGGTCCGCGACCAAGAGGCGC	300
QY	301	CTGGAACAAGATCGAGGAGGAGGACGAACAAGTCCAGGAGAAGACCGCAGGCCAAGAG	360
Db	301	CTGGAACAAGATCGAGGAGGAGGACGAACAAGTCCAGGAGAAGACCGCAGGCCAAGAG	360
QY	361	GCCGACGGCAAGGTGAAGCCAGAACTAATCCCATCTGTGGAACCTGCAGGGCCAGATGTTG	420
Db	361	GCCGACGGCAAGGTGAAGCCAGAACTAATCCCATCTGTGGAACCTGCAGGGCCAGATGTTG	420
QY	421	CACCAAGGCATCAGCCCCCGCAACCTTGAAACGCTTGGGTGAAGGTGATCGAGGAGAGGCC	480
Db	421	CACCAAGGCATCAGCCCCCGCAACCTTGAAACGCTTGGGTGAAGGTGATCGAGGAGAGGCC	480

QY	481	TTCAGCCCCGAGGTGATCCCATGTTTCAACCGCCCTGAGCGAGGGCGCCACCCCCAGAGAC	540
Db	481	TTTCAGCCCCGAGGTGATCCCATGTTTCAACCGCCCTGAGCGAGGGCGCCACCCCCAGAGAC	540
QY	541	CTGAACAACGATGTTGAACACCGTGGCGGCCACCAAGGCCGCATGCAAGTGTGAAGAC	600
Db	541	CTGAACAACGATGTTGAACACCGTGGCGGCCACCAAGGCCGCATGCAAGTGTGAAGAC	600
QY	601	ACCATCAACGAGAGAGGCCCGCCGAGTGGACCGCCTGCACCCCGTGCAAGCCGCCCCGTG	660
Db	601	ACCATCAACGAGAGAGGCCCGCCGAGTGGACCGCCTGCACCCCGTGCAAGCCGCCCCGTG	660
QY	661	GCCCCCGCCGATGCGCGACCCCCGCGCAGCGACATCGCCCGGCGCCACCAACCCCTG	720
Db	661	GCCCCCGCCGATGCGCGACCCCCGCGCAGCGACATCGCCCGGCGCCACCAACCCCTG	720
QY	721	CAGGAGCAGATCGCCTTGATGACCAAGAACCCCCCGTGCCCGTGGCGACATTTACAAG	780
Db	721	CAGGAGCAGATCGCCTTGATGACCAAGAACCCCCCGTGCCCGTGGCGACATTTACAAG	780
QY	781	CGGTGATCATCTCTGGGCTTGAAACAAGTCTGCGGATGTACAGCCCCGTGACATCCTG	840
Db	781	CGGTGATCATCTCTGGGCTTGAAACAAGTCTGCGGATGTACAGCCCCGTGACATCCTG	840
QY	841	GACATCCGCCAGGGCCCCCAAGAGACCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC	900
Db	841	GACATCCGCCAGGGCCCCCAAGAGACCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC	900
QY	901	CTGCGCGCCGAGCAGGCCACCCCAAGACGTGAAGAAGTGAATGACCGAGACCCTGCTGGTG	960
Db	901	CTGCGCGCCGAGCAGGCCACCCCAAGACGTGAAGAAGTGAATGACCGAGACCCTGCTGGTG	960
QY	961	CAGAACGCCAACCCCGACTGCAAGACCATCTCGCGCGCTCTCGGCCCCCGGCGCACCCCTG	1020
Db	961	CAGAACGCCAACCCCGACTGCAAGACCATCTCGCGCGCTCTCGGCCCCCGGCGCACCCCTG	1020
QY	1021	GAGGAGATGATGACCGCCTGCAAGGGCGGTGGGCGGCCCGGCCACAAGGCCCGCTGCTG	1080
Db	1021	GAGGAGATGATGACCGCCTGCAAGGGCGGTGGGCGGCCCGGCCACAAGGCCCGCTGCTG	1080
QY	1081	GCCGAGCGCATGAGCCAGGCCAACAGGTGAACATCATGATGACAGAAGCAACTTCAAG	1140
Db	1081	GCCGAGCGCATGAGCCAGGCCAACAGGTGAACATCATGATGACAGAAGCAACTTCAAG	1140
QY	1141	GCCCCCCCCGCGCAAGTCAAGTGTCTTCAACTGCGGCAAGAGGGCCACATCGCCAAAGAAC	1200
Db	1141	GCCCCCCCCGCGCGCAAGTCAAGTGTCTTCAACTGCGGCAAGAGGGCCACATCGCCAAAGAAC	1200
QY	1201	TGCGCGCGCCCCCGCAAGAGGGCTGTGAAATGCGGCAAGAGGGCCACCATGATGAG	1266
Db	1201	TGCGCGCGCCCCCGCAAGAGGGCTGTGAAATGCGGCAAGAGGGCCACCATGATGAG	1266
QY	1261	GACTGCACCGAGCGCCAGGCCCACTTCTTGCGGAAGATCTGCCCCAGCCACAAGGGCGCG	1326
Db	1261	GACTGCACCGAGCGCCAGGCCCACTTCTTGCGGAAGATCTGCCCCAGCCACAAGGGCGCG	1326
QY	1321	CCCGCAACTTCTCTGCAGAACCGCAGCGAGCCCGCGCCGCCACCGTGCCACCGCCCCC	1386
Db	1321	CCCGCAACTTCTCTGCAGAACCGCAGCGAGCCCGCGCCGCCACCGTGCCACCGCCCCC	1386
QY	1381	CCCGCCGAGAGCTTCCGCTTGAGAGAGACACCCCGCCCCCAAGCAGAGGCCCAAGGAC	1440
Db	1381	CCCGCCGAGAGCTTCCGCTTGAGAGAGACACCCCGCCCCCAAGCAGAGGCCCAAGGAC	1440
QY	1441	CGCGAGCCCTTACCGCGAGGCCCTTGACCGCCCTGCGCAGCCTGTTCGCGAGCGGCCCTG	1500
Db	1441	CGCGAGCCCTTACCGCGAGGCCCTTGACCGCCCTGCGCAGCCTGTTCGCGAGCGGCCCTG	1500
QY	1501	AGCCAGTAA 1509	
Db	1501	AGCCAGTAA 1509	

```
RESULT 4
ABL39958
ID ABL39958 standard; DNA; 1509 BP.
XX
AC ABL39958;
XX
DT 15-MAY-2002 (first entry)
XX
DE Synthetic Gag polynucleotide sequence SEQ ID NO:21.
XX
KW Human immunodeficiency virus type C; antigenic HIV type C protein;
KW immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
KW immunostimulant; gene therapy; gene; ds.
XX
OS Human immunodeficiency virus; type C.
XX
PN WO200204493-A2.
XX
PD 17-JAN-2002.
XX
PF 05-JUL-2001; 2001WO-US021241.
XX
PR 05-JUL-2000; 2000US-00610313.
XX
PA (CHIR ) CHIRON CORP.
PA (UYST-) UNIV STELLENBOSCH.
PI Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg EJ;
PI WPI; 2002-154920/20.
XX
DR
XX
PT New polynucleotides encoding antigenic HIV Type C polypeptides, useful in
PT applications including DNA immunization or generation of packaging cell
PT lines, particularly in gene therapy.
XX
PS Example 1; Fig 6; 233pp; English.
XX
CC The present invention describes expression cassettes comprising a
CC polynucleotide sequence encoding a polypeptide comprising immunogenic HIV
CC type C polypeptides. The expression cassettes comprise any of the HIV
CC type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or Nef
CC (1). (1) have immunostimulant activity and can be used in gene therapy.
CC The HIV type C polynucleotides are useful in applications including DNA
CC immunisation, generation of packaging cell lines, and production of HIV
CC type C proteins. The polynucleotides are particularly useful in gene
CC therapy and DNA immunisation applications. ABL39942 to ABL40054 and
CC ABB06204 to ABB06215 represent sequences used in the exemplification of
CC the present invention.
XX
SQ Sequence 1509 BP; 321 A; 560 C; 470 G; 158 T; 0 U; 0 Other;
Query Match 98.8%; Score 1491.4; DB 6; Length 1509;
Best Local Similarity 99.3%; Pred. No. 5.5e-175;
Matches 1498; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 1 ATGGGCGCCCGCGCCAGCATCTGCGCGCGAGAGCTGACCAAGTGGAGAGATCCGC 60
DB 1 ATGGGCGCCCGCGCCAGCATCTGCGCGCGAGAGCTGACCAAGTGGAGAGATCCGC 60
QY 61 CTGCGCCCGCGCGGCAAGAGCACTGCTGAAGCACTGTGTGGGCGAGCCGCGAG 120
DB 61 CTGCGCCCGCGCGGCAAGAGCACTGCTGAAGCACTGTGTGGGCGAGCCGCGAG 120
QY 121 CTGAGAGGCTTGCCTTGAACCCGCGCTGTGAGACCGCGAGGCTGCAAGCAATC 180
DB 121 CTGAGAGGCTTGCCTTGAACCCGCGCTGTGAGACCGCGAGGCTGCAAGCAATC 180
QY 181 ATGAAGCAGCTGACCGCGCTGAGACCGGACCGAGAGCTGCGACCTGTACAAC 240
DB 181 ATGAAGCAGCTGACCGCGCTGAGACCGGACCGAGAGCTGCGACCTGTACAAC 240
QY 241 ACCGTGGCCACCTGTACTGTGTCAGCGCGCATGAGGTCCGCGACCAAGAGAGGCC 300
```

```
DB 241 ACCGTGGCCACCTGTACTGTGTCAGCGCGCATGAGGTCCGCGACCAAGAGAGGCC 300
QY 301 CTGACCAAGATCGAGGAGCAGAACAACTCCAGCAGAAAGCCAGAGGCCAAGAGAG 360
DB 301 CTGACCAAGATCGAGGAGCAGAACAAAGCCAGCAGAAAGCCAGAGGCCAAGAGAG 360
QY 361 GCCGACGCAAGGTGAGCCAGAACTACCCATCGTGCAGAACTGCAAGGCCAGATGTG 420
DB 361 GCCGACGCAAGGTGAGCCAGAACTACCCATCGTGCAGAACTGCAAGGCCAGATGTG 420
QY 421 CACCAGGCCATCAGCCCGCACCCTGACGCTGGTGAAGTGAATCAGAGAGAGGCC 480
DB 421 CACCAGGCCATCAGCCCGCACCCTGACGCTGGTGAAGTGAATCAGAGAGAGGCC 480
QY 481 TTCAGCCCGAGGTGATCCCATGTTCAACCGCTGAGCGAGGGCGCCACCCCGAGGAC 540
DB 481 TTCAGCCCGAGGTGATCCCATGTTCAACCGCTGAGCGAGGGCGCCACCCCGAGGAC 540
QY 541 CTGAACACGATGTTGAACACGTTGAGCGGCGCACAGGCGCCATGCAAGTGTGAAGGAC 600
DB 541 CTGAACACGATGTTGAACACGTTGAGCGGCGCACAGGCGCCATGCAAGTGTGAAGGAC 600
QY 601 ACCATCAACGAGAGGCGCGAGTGGACCGCCTGACCCCGTGACGCGCGCCCGGTG 660
DB 601 ACCATCAACGAGAGGCGCGAGTGGACCGCCTGACCCCGTGACGCGCGCCCGGTG 660
QY 661 GCCCGCGCGCAGATGCGCGGACCCCGCGCGCAGCAGCATGCGCGCGCACAGCACCTG 720
DB 661 GCCCGCGCGCAGATGCGCGGACCCCGCGCGCAGCAGCATGCGCGCGCACAGCACCTG 720
QY 721 CAGAGCAGATGCTGATGACCAAGACCCCGCGTCCGCTGAGCGGCGCATCTACAAG 780
DB 721 CAGAGCAGATGCTGATGACCAAGACCCCGCGTCCGCTGAGCGGCGCATCTACAAG 780
QY 781 CGGTGATCATCTGGGCTGAACAAGATGTTGCGGATGTACAAGCCCGTGAGCATCTG 840
DB 781 CGGTGATCATCTGGGCTGAACAAGATGTTGCGGATGTACAAGCCCGTGAGCATCTG 840
QY 841 GACATCGCGCAGGCGCCCAAGAGGCGCTTCGCGACTAGCTGAGCGCTTCAAGACC 900
DB 841 GACATCGCGCAGGCGCCCAAGAGGCGCTTCGCGACTAGCTGAGCGCTTCAAGACC 900
QY 901 CTGCGCGCGAGCAGGCGCACCGAGAGCTGAAGAACTGATGACCGGACCTGTGCTG 960
DB 901 CTGCGCGCGAGCAGGCGCACCGAGAGCTGAAGAACTGATGACCGGACCTGTGCTG 960
QY 961 CAGAACGCCAACCCCGACTGCAAGACATCTGCGGCTCTCGGCGCCGCGCACCTG 1020
DB 961 CAGAACGCCAACCCCGACTGCAAGACATCTGCGGCTCTCGGCGCCGCGCACCTG 1020
QY 1021 GAGGAGATGATGACCGCTGCGAGGCGCTGGGCGCGCCCGCCACAAGCGCGCTGCTG 1080
DB 1021 GAGGAGATGATGACCGCTGCGAGGCGCTGGGCGCGCCCGCCACAAGCGCGCTGCTG 1080
QY 1081 GCCGAGCGATGAGCGCCACAGCGTGAACATCATGATGCAAGAGCAACTTCAAG 1140
DB 1081 GCCGAGCGATGAGCGCCACAGCGTGAACATCATGATGCAAGAGCAACTTCAAG 1140
QY 1141 GCGCCCGCGCGCAACGTCATGCTTCAACTGCGGCAAGAGAGGCCCAATCGCCAAGAC 1200
DB 1141 GCGCCCGCGCGCAACGTCATGCTTCAACTGCGGCAAGAGAGGCCCAATCGCCAAGAC 1200
QY 1201 TGCCGCGCGCGCGCAAGAGGCTGCTGGAAGTGCAGAGAGGCGCCACAGATGAAG 1260
DB 1201 TGCCGCGCGCGCGCAAGAGGCTGCTGGAAGTGCAGAGAGGCGCCACAGATGAAG 1260
QY 1261 GACTGCAACGAGCGCGCGCAACTTCTGGGCAAGATCTGGCCAGGCCCAAGGGCGCG 1320
DB 1261 GACTGCAACGAGCGCGCGCAACTTCTGGGCAAGATCTGGCCAGGCCCAAGGGCGCG 1320
QY 1321 CCGGCAACTTCTGTGCAAGACCGCAGAGCGCGCGCGCCCAACCGTGGCCACCGCGCCC 1380
DB 1321 CCGGCAACTTCTGTGCAAGACCGCAGAGCGCGCGCGCCCAACCGTGGCCACCGCGCCC 1380
```







DB 1141 GGGCCCCCGCGCAAGTGAAGTCTTCACTGCGGCAAGGAGGGCCACATCGCCAAAGAAC 1200  
QY 1201 TGCCGCGCCCCCGCAAGAGGCTGTGGAAGTGGCGCAAGAGGGCCACCAAGATGAAG 1260  
DB 1201 TGCCGCGCCCCCGCAAGAGGCTGTGGAAGTGGCGCAAGAGGGCCACCAAGATGAAG 1260  
QY 1261 GACTGCAACGAGCGCCAGGCCAATCTCTGGGCAAGATCTGGCCAGCCACAAGGGCGCG 1320  
DB 1261 GACTGCAACGAGCGCCAGGCCAATCTCTGGGCAAGATCTGGCCAGCCACAAGGGCGCG 1320  
QY 1321 CCGGCAACTTCTCTGCAAGACCGAGCGCGCGCCGCCCAAGTGGCCCAAGCGCCCGCC 1380  
DB 1321 CCGGCAACTTCTCTGCAAGACCGAGCGCGCGCGCCGCCCAAGTGGCCCAAGCGCCCGCC 1380  
QY 1381 CCGGCGAGAGCTTCTGCGAGAGAGCAACCGCGCGCGCCGCCCAAGAGCGCGCAAGGAC 1440  
DB 1381 CCGGCGAGAGCTTCTGCGAGAGAGCAACCGCGCGCGCGCCGCCCAAGAGCGCGCAAGGAC 1440  
QY 1441 CCGGAGCCCTAAGCGAGAGCGCGCTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1500  
DB 1441 CCGGAGCCCTAAGCGAGAGCGCGCTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1500  
QY 1501 AGCCAGTAA 1509  
DB 1501 AGCCAGTAA 1509

RESULT 6  
AAL44553  
ID AAL44553 standard; DNA; 1509 BP.

AC AAL44553;  
XX

DT 29-AUG-2003 (revised)  
DT 08-NOV-2002 (first entry)

XX HIV-1 p55gag polypeptide coding sequence 4.

XX HIV; ds: vaccine; gene; immune response; microparticle;  
KW adsorbent surface; poly(alpha-hydroxy acid); polyhydroxy butyric acid;  
KW polycaprolactone; polyorthoester; polycyanoacrylate; detergent;  
KW submicron emulsion; viral infection; bacterial infection;  
KW parasitic infection; HIV-1 p55gag polypeptide.

XX Human immunodeficiency virus 1.

XX Key Location/Qualifiers

FT variation replace(282, C)

FT variation replace(331, T)

FT variation replace(332, C)

FT variation replace(549, G)

FT variation replace(553, T)

FT variation replace(783, G)

FT variation replace(816, G)

FT variation replace(999, T)

FT variation replace(1002, C)

FT variation replace(1089, G)

FT variation replace(1149, G)

FT variation replace(1158, C)

FT variation replace(1158, C)

FT variation replace(1158, C)

FT variation replace(1158, C)

FT variation replace(1158, C)

FT variation replace(1158, C)

FT variation replace(1158, C)

FT variation replace(1158, C)

FT variation replace(1158, C)

FT variation replace(1158, C)

FT variation replace(1158, C)

FT variation replace(1158, C)

FT variation replace(1158, C)

FT variation replace(1158, C)

FT variation replace(1158, C)

FT variation replace(1158, C)

XX 04-APR-2002.  
PD 28-SEP-2001; 2001WO-US030540.  
XX 28-SEP-2001; 2001WO-US030540.  
PF 28-SEP-2001; 2001WO-US030540.  
XX 28-SEP-2001; 2001WO-US030540.  
PR 28-SEP-2001; 2001WO-US030540.  
XX 28-SEP-2001; 2001WO-US030540.  
PR 28-SEP-2001; 2001WO-US030540.  
XX 28-SEP-2001; 2001WO-US030540.  
PA (CHIR ) CHIRON CORP.  
XX O'hagan D, Otten G, Donnelly JJ, Polo JM, Barnett S, Singh M;  
PI Ulmer J, Dubensky TW;  
XX  
XX MPI; 2002-519084/55.  
DR  
XX  
XX A microparticle to which a biologically active macromolecule is adsorbed,  
PT for use as a vaccine composition to treat viral, bacterial or parasitic  
PT infections, comprises a polymer microparticle, a detergent and a  
PT submicron emulsion.  
XX  
XX Claim 72, Fig 6; 100pp; English.  
XX  
XX The invention relates to a method of raising an immune response in a host  
CC animal. The method of the invention comprises administering a  
CC microparticle that has an adsorbent surface to which a first biologically  
CC active macromolecule (e.g. a nucleic acid) has been adsorbed. The  
CC microparticle comprises a polymer microparticle of poly(alpha-hydroxy  
CC acid), a polyhydroxy butyric acid, a polycaprolactone, a polyorthoester,  
CC a polycyanoacrylate, a detergent, and submicron emulsion. The method/  
CC microparticle of the invention is useful for immunising a host animal  
CC against viral, bacterial or parasitic infections. The present DNA  
CC sequence encodes a HIV-1 p55gag polypeptide. (Updated on 29-AUG-2003 to  
CC standardise OS field)  
XX  
SQ Sequence 1509 BP; 321 A; 559 C; 471 G; 158 T; 0 U; 0 Other;

Query Match 98.7%; Score 1489.8; DB 6; Length 1509;  
Best Local Similarity 99.2%; Pred. No. 8.6e-175;  
Matches 1497; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 ATGGGCGCCCGCGCGCAAGCACTACATCTGAAGCACTGTGTGGCGCCGCGAG 120  
DB 61 CTGGCGCCCGCGCGCGCAAGCACTACATCTGAAGCACTGTGTGGCGCCGCGAG 120  
QY 121 CTGGAGGGCTTGGCTTGAACCCCGGCTGTGAGACCGCGGCTGCAAGCAGATC 180  
DB 121 CTGGAGGGCTTGGCTTGAACCCCGGCTGTGAGACCGCGGCTGCAAGCAGATC 180  
QY 181 ATGAAGCAGCTGCAAGCCCGCTGTGAGACCGCGGCTGTGAGACCGCGGCTGTGAGAC 240  
DB 181 ATGAAGCAGCTGCAAGCCCGCTGTGAGACCGCGGCTGTGAGACCGCGGCTGTGAGAC 240  
QY 241 ACCGTGGCAACCTGTACTGTGTCAGCGCGGCTGTGAGACCGCGGCTGTGAGACCGCGG 300  
DB 241 ACCGTGGCAACCTGTACTGTGTCAGCGCGGCTGTGAGACCGCGGCTGTGAGACCGCGG 300  
QY 301 CTGACACAGATCGAGAGAGCAGAACAGTCCAGCAGAACCCAGCAGGCCAAGAGAG 360  
DB 301 CTGACACAGATCGAGAGAGCAGAACAGTCCAGCAGAACCCAGCAGGCCAAGAGAG 360  
QY 361 GCGGAGCGGCAAGGTGAGCGCAAGTACCCCATCTGTGCAAGACCTGCAAGGCCAAGTGTG 420  
DB 361 GCGGAGCGGCAAGGTGAGCGCAAGTACCCCATCTGTGCAAGACCTGCAAGGCCAAGTGTG 420  
QY 421 CACCAAGGCATCAGCCCGCGCAACCTGTGAGACCGCGGCTGTGAGAGGTGATCGAGAGAGGCC 480  
DB 421 CACCAAGGCATCAGCCCGCGCAACCTGTGAGACCGCGGCTGTGAGAGGTGATCGAGAGAGGCC 480  
QY 481 TTCAGCCCGGAGGTGATCCCATCTGTGACCGCGCTGTGAGAGAGGCCAAGGCCAAGAGAC 540

```
DB 481 TTCAAGCCCGAGGTGATCCCATGTTTCAACGCGCTGAGCGAGGGCGCACCCCCCAGAGAC 540
QY 541 CTGAACACGATGTTGAACACCGTGGGGCGGCAACAGCGCCCATGAGATGCTGAAGGAC 600
DB 541 CTGAACACGATGCTGAACACCGTGGGGCGGCAACAGCGCCCATGAGATGCTGAAGGAC 600
QY 601 ACCATCAACGAGGAGGCGCGCGAGTGGGACCGCCTGCAACCCCTGCAAGCGCGCCCGCTG 660
DB 601 ACCATCAACGAGGAGGCGCGCGAGTGGGACCGCCTGCAACCCCTGCAAGCGCGCCCGCTG 660
QY 661 GCGCGCGCGCGAGATGCGCGGACCGCGCGGAGACATCGCGCGGCGGCAACGAGACCGCTG 720
DB 661 GCGCGCGCGCGAGATGCGCGGACCGCGCGGAGACATCGCGCGGCGGCAACGAGACCGCTG 720
QY 721 CAGGAGCAGATCGCCTGATGATGACGCAACCCCGCTGCGCGGCGGAGATCTACAAG 780
DB 721 CAGGAGCAGATCGCCTGATGATGACGCAACCCCGCTGCGCGGCGGAGATCTACAAG 780
QY 781 CGGTGATGATCTCTGGGCTGAAACAAGATCGTGGATGTAACGCGCGCTTCTTCAAGACC 840
DB 781 CGGTGATGATCTCTGGGCTGAAACAAGATCGTGGATGTAACGCGCGCTTCTTCAAGACC 840
QY 841 GACATCGCGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
DB 841 GACATCGCGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
QY 901 CTGCGCGCGCGAGCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 960
DB 901 CTGCGCGCGCGAGCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 960
QY 961 CAGAACCGCAACCCCGACTGCAAGACCATCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 1020
DB 961 CAGAACCGCAACCCCGACTGCAAGACCATCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 1020
QY 1021 GAGGAGATGATGACCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1080
DB 1021 GAGGAGATGATGACCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1080
QY 1081 GCGGAGGCGATGAGCGCGCGCAACAGCGTGAACATCATGATGCAAGAGCAACTTCAAG 1140
DB 1081 GCGGAGGCGATGAGCGCGCGCAACAGCGTGAACATCATGATGCAAGAGCAACTTCAAG 1140
QY 1141 GCGCGCGCGCGCGCAACAGTGTCTTCAACTGCGCGCGCGCGCGCGCGCGCGCGCGCG 1200
DB 1141 GCGCGCGCGCGCGCAACAGTGTCTTCAACTGCGCGCGCGCGCGCGCGCGCGCGCGCG 1200
QY 1201 TGCGCGCGCGCGCGCGCAAGAGGCGCTGCTGAAGTGCAGGCAAGAGGCGCGCGCG 1260
DB 1201 TGCGCGCGCGCGCGCGCAAGAGGCGCTGCTGAAGTGCAGGCAAGAGGCGCGCGCG 1260
QY 1261 GACTGCAACCGAGCGCGCGCGCAACTTCTGCGCGCAAGATCTGCGCGCAAGAGCGCG 1320
DB 1261 GACTGCAACCGAGCGCGCGCGCAACTTCTGCGCGCAAGATCTGCGCGCAAGAGCGCG 1320
QY 1321 CCGCGCAACTTCTCTGCAAGAACCGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
DB 1321 CCGCGCAACTTCTCTGCAAGAACCGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
QY 1381 CCGCGCGAGAGCTTCTGCGCTTGAAGAGACCAACCGCGCGCGCGCGCGCGCGCGCG 1440
DB 1381 CCGCGCGAGAGCTTCTGCGCTTGAAGAGACCAACCGCGCGCGCGCGCGCGCGCGCG 1440
QY 1441 CCGGAGCGCTTACCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1500
DB 1441 CCGGAGCGCTTACCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1500
QY 1501 AGCGAGTAA 1509
DB 1501 AGCGAGTAA 1509
```

RESULT 7

```
ABL39972
ID ABL39972 standard; DNA; 1494 BP.
XX
AC ABL39972;
XX
DT 15-MAY-2002 (first entry)
XX
DB Synthetic Gag polynucleotide sequence SEQ ID NO:51.
XX
KW Human immunodeficiency virus type C; antigenic HIV type C protein;
KW immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
KW immunostimulant; gene therapy; gene; ds.
XX
OS Human immunodeficiency virus; type C.
OS Synthetic.
XX
FN WO200204493-A2.
XX
PD 17-JAN-2002.
XX
PF 05-JUL-2001; 2001WO-US021241.
XX
PR 05-JUL-2000; 2000US-00610313.
XX
PA (CHIR ) CHIRON CORP.
PA (UYST-) UNIV STELLENBOSCH.
XX
PI Zur Megeide J, Barnett SW, Engelbrecht S, Van Rensburg EJ;
XX
DR WPI; 2002-154920/20.
XX
PT New polynucleotides encoding antigenic HIV Type C polypeptides, useful in
PT applications including DNA immunization or generation of packaging cell
PT lines, particularly in gene therapy.
XX
PS Claim 7; Fig 22; 233pp; English.
XX
CC The present invention describes expression cassettes comprising a
CC polynucleotide sequence encoding a polypeptide comprising immunogenic HIV
CC type C polypeptides. The expression cassettes comprise any of the HIV
CC type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or Nef
CC (I). (I) have immunostimulant activity and can be used in gene therapy.
CC The HIV type C polynucleotides are useful in applications including DNA
CC immunisation, generation of packaging cell lines, and production of HIV
CC Type C proteins. The polynucleotides are particularly useful in gene
CC therapy and DNA immunisation applications. ABL39942 to ABL40054 and
CC ABB06204 to ABB06215 represent sequences used in the exemplification of
CC the present invention
XX
SQ Sequence 1494 BP; 325 A; 557 C; 455 G; 157 T; 0 U; 0 Other;
Query Match 86.0%; Score 1297.6; DB 6; Length 1494;
Best Local Similarity 92.7%; Pred. No. 3.3e-151;
Matches 1402; Conservative 0; Mismatches 89; Indels 21; Gaps 3;
QY 1 ATGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
DB 1 ATGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
QY 61 CTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
DB 61 CTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
QY 121 CTGAGAGGCGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
DB 121 CTGAGAGGCGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
QY 181 ATGAAGCAGCTGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
DB 181 ATCAAGCAGCTGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
QY 241 ACCGTGGCCACCTGTACTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
DB 241 ACCGTGGCCACCTGTACTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
```

```

Db 241 ACCGTGGCCACCTGTACTGCTGTCACAAAGGGCATCGAGGTGCGGACACCAAGGAGGCC 300
QY 301 CTGGACAAGATCGAGGAGAGCAGAAACAAGTCCCAAGAGAACCAGCAGGCCAAGGAG 360
Db 301 CTGGACAAGATCGAGGAGAGCAGAAACAAGTCCCAAGAGAACCAGCAGGCCAAGGAGGCC 360
QY 361 GCCGACGGCAAGGTGAGCCAGAACTAACCCTATGTCGAGAACTTGACGGGCCAGATGTC 420
Db 361 GCCGACGGCAAGGTGAGCCAGAACTAACCCTATGTCGAGAACTTGACGGGCCAGATGTC 420
QY 421 CACCAAGGCATCAGCCCCCGCACCCTGAAACGCTGGTGAAGGTGATCGAGAGAAAGGCC 480
Db 421 CACCAAGGCATCAGCCCCCGCACCCTGAAACGCTGGTGAAGGTGATCGAGAGAAAGGCC 480
QY 481 TTCAGCCCCGAGGTGATCCCATGTTCAACCGCCCTGAGCGAGGGCGCCACCCCGCAGGAC 540
Db 481 TTCAGCCCCGAGGAGATCCCATGTTCAACCGCCCTGAGCGAGGGCGCCACCCCGCAGGAC 540
QY 541 CTGAACACGATGTTGAACACCGTGGGCGGCCACAGGCCCGCATGCAAGATCTGAAGGAC 600
Db 541 CTGAACACCATGTGTAACACCGTGGGCGGCCACAGGCCCGCATGCAAGATCTGAAGGAC 600
QY 601 ACCATCAACGAGGAGGCCCGCGAGTGGAGCCGCTGCAACCCCGTGCAGGCCGCGCCGTG 660
Db 601 ACCATCAACGAGGAGGCCCGCGAGTGGAGCCGCTGCAACCCCGTGCAGGCCGCGCCGTG 660
QY 661 GCCCGCCGCGAGATGCGCGAACCCCGCGCGCAGCAATCGCGCGGCCACCAACCCCTG 720
Db 661 GCCCGCCGCGAGATGCGCGAACCCCGCGCGCAGCAATCGCGCGGCCACCAACCCCTG 720
QY 721 CAGGACGAGATCGCTGATGACAGCAACCCCGCTGCGCCGTGGGCGACATCTAACAG 780
Db 721 CAGGACGAGATCGCTGATGACAGCAACCCCGCTGCGCCGTGGAGACATCTAACAG 780
QY 781 GGGTGAATCATCTTGAGCTGAAACAAGATCGTGGGATGTACAGCCCGTGAACATCTCTG 840
Db 781 GGGTGAATCATCTTGAGCTGAAACAAGATCGTGGGATGTACAGCCCGTGAACATCTCTG 840
QY 841 GACATCCGCGCAGGCGCCCAAGAGAGCCCTTCGCGACTACGTGGAACCGCTTCTTCAAGACC 900
Db 841 GACATCAAGCAGGCGCCCAAGAGAGCCCTTCGCGACTACGTGGAACCGCTTCTTCAAGACC 900
QY 901 CTGCGCGCCGAGGAGGCCACCCAGAGCGTGAAGAACTGATGACCGAGACCTGCTGTG 960
Db 901 CTGCGCGCCGAGGAGGCCACCCAGAGCGTGAAGAACTGATGACCGAGACCTGCTGTG 960
QY 961 CAGAAACGCCAACCCCGACTGCAAGAACCATCTGCGCGCTCTGCGCCCGCGCCACCTG 1020
Db 961 CAGAAACGCCAACCCCGACTGCAAGAACCATCTGCGCGCTCTGCGCCCGCGCCACCTG 1020
QY 1021 GAGGAGATGATGACCGCTGCGCAGGCGGTGCGCGGCCCGCCACAAGGCCCGCTGCTG 1080
Db 1021 GAGGAGATGATGACCGCTGCGCAGGCGGTGCGCGGCCCGCCACAAGGCCCGCTGCTG 1080
QY 1081 GCCGAGGCGATGAGCCAGGCCAACAGCGTGAACATCATGATGCAAGAGCAACTTCAAG 1140
Db 1081 GCCGAGGCGATGAGCCAGGCCAACAGCGTGAACATCATGATGCAAGAGCAACTTCAAG 1140
QY 1141 GGGCCCCCGGCGCAAGTCAAGTCTTCAACTGCGGCAAGGAGGGCCACATGCGCAAGAAC 1200
Db 1138 GGGCAGCAACCGCATCATCAAGTCTTCAACTGCGGCAAGGAGGGCCACATGCGCGCAAC 1197
QY 1201 TGCCCGCGCCCCCGCAAGAGGGCTGCTGGAAGTGGCGCAAGAGGGCCACAGATGAAG 1260
Db 1198 TGCCCGCGCCCCCGCAAGAGGGCTGCTGGAAGTGGCGCAAGAGGGCCACAGATGAAG 1257
QY 1261 GACTGCAACGAGCGCCAGGCCAACTTCTTGGGCAAGATTCGCCCAAGCAAGGGCGGC 1320
Db 1258 GACTGCAACGAGCGCCAGGCCAACTTCTTGGGCAAGATTCGCCCAAGCAAGGGCGGC 1317
QY 1321 CCGCGCAACTTCTGCAAGAACCGCAGCGAGC---CCGCGCGCCCCCACCCTGCCACCGGCC 1377
Db 1318 CCGCGCAACTTCTGCAAGAACCGCGCGAGGCCACCGCGCCCCCGCGCGAGCCACCGGCC 1377

```

```

QY 1378 CCCCCCGCGAGAGCTTCGCTTCGAGAGACCAACCCCGCCCCCAAGCAGAGCCCAAG 1437
Db 1378 CCCCCCGCGAGAGCTTCGCTTCGAGAGACCAACCCCGCCCCCGCAAGGAGAGAG 1437
QY 1438 GACCGCGAGCCCTTACCGAGAGCCCTGACCGCCCTGCGCAGCCTGTTGCGCAGCGGCC 1497
Db 1438 -----CGCAGCCCTGACCGAGCCTGAAGAGCCTGTTGCGCAGCGAGCCCC 1482
QY 1498 CTGAGCCAGTTAA 1509
Db 1483 CTGAGCCAGTTAA 1494

RESULT 8
ABLA0020
ID ABL40020 standard; DNA; 1491 BP.
XX
AC ABL40020;
XX
DT 15-MAY-2002 (first entry)
XX
DE Synthetic Gag polynucleotide sequence SEQ ID NO:99.
XX
KW Human immunodeficiency virus type C; antigenic HIV type C protein;
KW immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
KW immunostimulant; gene therapy; gene; ds.
XX
OS Human immunodeficiency virus; type C.
XX
PN WO200204493-A2.
XX
PD 17-JAN-2002.
XX
PF 05-JUL-2001; 2001WO-US021241.
XX
PR 05-JUL-2000; 2000US-00610313.
XX
PA (CHIR ) CHIRON CORP.
PA (UYST-) UNIV STELLENBOSCH.
XX
PI Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg EJ;
XX
DR WPI; 2002-154920/20.
XX
PT New polynucleotides encoding antigenic HIV Type C polypeptides, useful in
PT applications including DNA immunization or generation of packaging cell
PT lines, particularly in gene therapy.
XX
PS Claim 8; Fig 70; 233pp; English.
XX
CC The present invention describes expression cassettes comprising a
CC polynucleotide sequence encoding a polypeptide comprising immunogenic HIV
CC type C polypeptides. The expression cassettes comprise any of the HIV
CC type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or Nef
CC (I). (I) have immunostimulant activity and can be used in gene therapy.
CC The HIV type C polynucleotides are useful in applications including DNA
CC immunisation, generation of packaging cell lines, and production of HIV
CC type C proteins. The polynucleotides are particularly useful in gene
CC therapy and DNA immunisation applications. ABL39942 to ABL40054 and
CC ABB06204 to ABB06215 represent sequences used in the exemplification of
CC the present invention
XX
SQ Sequence 1491 BP; 319 A; 557 C; 457 G; 158 T; 0 U; 0 Other;

Query Match 85.8%; Score 1294.8; DB 6; Length 1491;
Best Local Similarity 92.7%; Pred. No. 7.3e-151;
Matches 1401; Conservative 0; Mismatches 87; Indels 24; Gaps 3;

QY 1 ATGGGCGCGCGCGCCAGCATCTTGGCGCGGAGAGCTGACCAAGTGGAAGATCCGC 60
Db 1 ATGGGCGCGCGCGCCAGCATCTTGGCGCGGAGAGCTGACCAAGTGGAAGATCCGC 60

```

```

QY 61 CTGCGCCCCGCGCAAGAGCACTACATGCTGAAGCACTGCTGTGTGGCCAGCCGCGAG 120
DB 61 CTGCGCCCCGCGCGCGCAAGAGCACTACATGCTGAAGCACTGCTGTGTGGCCAGCCGCGAG 120
QY 121 CTGAGAGGCTTCCGCTGTAACCCCGGCTGTGTGAAGACCGCGAGGGCTGCAAGCATC 180
DB 121 CTGAGAGGCTTCCGCTGTAACCCCGGCTGTGTGAAGACCGCGAGGGCTGCAAGCATC 180
QY 181 ATGAAGCAGCTGCAAGCCCGCTGCAAGCCCGCATCGAGAGCTGCGCAGCTGTACAAC 240
DB 181 ATGAAGCAGCTGCAAGCCCGCTGCAAGCCCGCATCGAGAGCTGCGCAGCTGTACAAC 240
QY 241 ACCGTGCGCACTCTGTACTGCTGTGACGCGCGCATCGAGTCCGCGACCAAGAGAGCC 300
DB 241 ACCGTGCGCACTCTGTACTGCTGTGACGCGCGCATCGAGTCCGCGACCAAGAGAGCC 300
QY 301 CTGGAACAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
DB 301 CTGGAACAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
QY 361 GCGGACGCGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
DB 361 GCGGACGCGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 421 CACCAAGGCGCATGAGCCCGCGCACTGTAAGCGCTGGGTGAAGGTGATCGAGAGAGAGAG 480
DB 421 CACCAAGGCGCATGAGCCCGCGCACTGTAAGCGCTGGGTGAAGGTGATCGAGAGAGAGAG 480
QY 481 TTCAAGCCCCGAGGTGATCCCAATGTTCAACCGCTGAGCGAGAGAGAGAGAGAGAGAG 540
DB 481 TTCAAGCCCCGAGGTGATCCCAATGTTCAACCGCTGAGCGAGAGAGAGAGAGAGAGAG 540
QY 541 CTGAACAACGATGTTGAACAACCGTGGGCGCGCAAGCGCGCGCATGCAAGATGCTGAAGAG 600
DB 541 CTGAACAACGATGTTGAACAACCGTGGGCGCGCAAGCGCGCGCATGCAAGATGCTGAAGAG 600
QY 601 ACCATCAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
DB 601 ACCATCAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY 661 GCGGCGCGCGAGATGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
DB 661 GCGGCGCGCGAGATGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 721 CAGGAGCAGATCGCTGTGATGACAGCAACCCCGCTGCGCGCGAGAGAGAGAGAGAGAG 780
DB 721 CAGGAGCAGATCGCTGTGATGACAGCAACCCCGCTGCGCGCGAGAGAGAGAGAGAGAG 780
QY 781 CGGTGATCATCTCTGTGGGCTGTAACAAGATGCTGCGCATGTACAGCCCGTGAAGCATCTG 840
DB 781 CGGTGATCATCTCTGTGGGCTGTAACAAGATGCTGCGCATGTACAGCCCGTGAAGCATCTG 840
QY 841 GACATCCCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
DB 841 GACATCCCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
QY 901 CTGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
DB 901 CTGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
QY 961 CAGAACGCGCAACCCCGACTGCAAGACCATCTGCGCGCTGCGCGCGCGCGCGAGAGAGAG 1020
DB 961 CAGAACGCGCAACCCCGACTGCAAGACCATCTGCGCGCTGCGCGCGCGCGCGAGAGAGAG 1020
QY 1021 GAGGAGATGATGACCGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
DB 1021 GAGGAGATGATGACCGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
QY 1081 GCGGAGGCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
DB 1081 GCGGAGGCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140

```

```

QY 1141 GCGGCGCGCGCAAGTCAAGTCTTCAACTGCGGCAAGAGAGAGAGAGAGAGAGAGAGAG 1200
DB 1141 GCGGCGCGCGCGCGCGCGCGTGAAGTCTTCAACTGCGGCGCGAGAGAGAGAGAGAGAGAG 1200
QY 1201 TGCGCGCGCGCGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
DB 1201 TGCGCGCGCGCGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
QY 1261 GACTGCACCGAGCGCGCAAGCACTTCTGGGCAAGATCTGCGCGAGAGAGAGAGAGAGAGAG 1320
DB 1261 GACTGCACCGAGCGCGCAAGCACTTCTGGGCAAGATCTGCGCGAGAGAGAGAGAGAGAGAG 1320
QY 1321 CCGGCGCACTTCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1377
DB 1321 CCGGCGCACTTCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1377
QY 1378 CCGGCGCGCGAGAGCTTCCGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1437
DB 1381 CCGGCGCGCGAGAGCTTCAAGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1431
QY 1438 GACCGCGAGAGCTTACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1497
DB 1432 G-----ACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1479
QY 1498 CTGAGCCAGTAA 1509
DB 1480 CTGAGCCAGTAA 1491

```

RESULT 9  
AAAS1609  
ID AAAS1609 standard; DNA; 1479 BP.

```

XX AC AAAS1609;
XX DT 31-OCT-2000 (first entry)
XX DB HIV synthetic Gag polynucleotide.
XX KW Gag; expression cassette; antigenic; type C; HIV; Env; synthetic;
XX KW DNA immunization; packaging cell line; antigen presentation; ss.
XX OS Human immunodeficiency virus; type C strain AF110965.
XX FH Key Location/Qualifiers
XX FT CDS 1..1479
XX FT /product= "Synthetic Gag"
XX FT /note= "Codon usage pattern was modified and inhibitory
XX FT elements (INS) and RRE sites were inactivated resulting
XX FT in improved expression"
XX PN MO200039304-A2.
XX PD 06-JUL-2000.
XX PF 30-DEC-1999; 99WO-US031273.
XX PR 31-DEC-1998; 98US-0114495P.
XX PR 01-SEP-1999; 99US-0152195P.
XX PA (CHIR ) CHIRON CORP.
XX PI Barnett S, Zur Megede J;
XX DR WPI; 2000-452401/39.
XX DR P-PSDB; AAY96943.
XX PT Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV Env
XX PT polypeptide and the polypeptide useful for immunizing a mammal especially
XX PT human against HIV.

```



PS Claim 2; Page 92-93; 113pp; English.

XX Expression cassettes comprising a polynucleotide encoding antigenic type  
CC C human immunodeficiency virus (HIV) Gag or Env polypeptides are useful  
CC in DNA immunization, generation of packaging cell lines and production of  
CC Gag- and/or Env-containing proteins. Synthetic Env and Gag expression  
CC cassettes exhibit increased potency for induction of cytotoxic T-  
CC lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1 self-  
CC assemble into non-infectious virus-like particles which are used as a  
CC matrix for the proper presentation of an antigen entrapped or associated  
CC to the immune system of the host

XX Sequence 1479 BP; 325 A; 529 C; 463 G; 162 T; 0 U; 0 Other;

Query Match 84.6%; Score 1276.8; DB 3; Length 1479;  
Best Local Similarity 92.5%; Pred. No. 1.2e-148;  
Matches 1399; Conservative 0; Mismatches 77; Indels 36; Gaps 4;

```

OY 1 ATGGCGCGCGCGCGCGCATCTGCGCGCGAGAGCTGAGCAAGTGGAGAAATCCGC 60
DB 1 ATGGCGCGCGCGCGCGCATCTGCGCGCGAGAGCTGAGCAAGTGGAGAAATCCGC 60
OY 61 CTGCGCGCGCGCGCGCGAGAGCACTACATGCTGAGACACTGTGTGGCCAGCCGCGAG 120
DB 61 CTGCGCGCGCGCGCGCGAGAGAGTGTACTATGATGAGCACTGTGTGTGGCCAGCCGCGAG 120
OY 121 CTGAGAGGCTTCCGCTGAAACCCCGCTGTGAGACCGCGAGGCTGCAAGCAGATC 180
DB 121 CTGAGAGAGTCCGCTTGAACCCCGCTGTGAGACCGCGAGGCTGCAAGCAGATC 180
OY 181 ATGAAGAGCTGACCGCGCGCTGACAGACCGGCAACGAGAGCTGCGCAGCTGTACAC 240
DB 181 ATCGCGCAGCTGACCGCGCGCTGACAGACCGGCAACGAGAGCTGTGTCAAC 240
OY 241 ACCGTGGCCACCTCTGTACTGCGTGACCGCGGCAATCGAGTCCGCGACCAAGAGGCC 300
DB 241 ACCGTGGCCACCTCTGTACTGCGTGACAGAGAGATCGAGTCCGCGACCAAGAGGCC 300
OY 301 CTGCAACAAGATCGAGAGAGAGAGAGCAAGTCCAGAGAGAGAGAGAGAGAGAGAG 360
DB 301 CTGCAACAAGATCGAGAGAGAGAGAGAGCAAGTCCAGAGAGAGATCCAGAGAGAGAGAG 360
OY 361 GCGCAGC---GGCAAGGTAGAGCAGAACTAACCCCATCTGAGAGAGAGAGAGAGAG 417
DB 361 GCGCAGCAGGCGAAGGTAGAGCAGAACTAACCCCATCTGAGAGAGAGAGAGAGAGAG 420
OY 418 GTGACACAGGCGCATAGACCCCGCAGACCTGAAACGCTGAGGTGAAAGGTATGAGAGAG 477
DB 421 GTGACACAGGCGCATAGACCCCGCAGACCTGAAACGCTGAGGTGAAAGGTATGAGAGAG 480
OY 478 GCGTTCAAGCCCGAGGTATCCCATGTTCAACCGCCTGAGCGAGAGGCGCCACCCCGCAG 537
DB 481 GCGTTCAAGCCCGAGGTATCCCATGTTCAACCGCCTGAGCGAGAGGCGCCACCCCGCAG 540
OY 538 GACCTGAACACGATGTTGAACACCGTGGGCGGCGCAACGAGCGCGCATGCAATGCTGAAG 597
DB 541 GACCTGAACACGATGTTGAACACCGTGGGCGGCGCAACGAGCGCGCATGCAATGCTGAAG 600
OY 598 GACACCATCAACGAGAGAGCGCGCGAGTGGAGCGCGCTGACACCCCGTGCAGGCGCGGCC 657
DB 601 GACACCATCAACGAGAGAGCGCGCGAGTGGAGCGCGCTGACACCCCGTGCAGGCGGCC 660
OY 658 GTGGCCCCCGGCGAGAGCGGAGACCCCGCGGAGAGAGATGCGCGCGCGCGCGAGAGAG 717
DB 661 ATGCGCCCCCGGCGAGAGCGGAGAGCGCGCGGAGAGATGCGCGCGCGCGCGAGAGAG 720
OY 718 CTGCAAGAGAGAGATGCGCTGATGAGCAAGCAACCCCGTGCAGGCGCGAGAGAGATG 777
DB 721 CTGCAAGAGAGAGATGCGCTGATGAGCAAGCAACCCCGTGCAGGCGCGAGAGATG 780
OY 778 AAGCGGTGATCATCTGGGCGCTGAACAAGATGCTGCGGATGTACAGCCCGTGAAGATC 837
DB 781 AAGCGGTGATCATCTGGGCGCTGAACAAGATGCTGCGGATGTACAGCCCGTGAAGATC 840

```

```

OY 838 CTGACATCCGCGAGGCGCCCAAGAGAGCCCTTCCGCGACTACGTGAGCCCTTCTTCAAG 897
DB 841 CTGACATCAAGCAGGCGCCCAAGAGAGCCCTTCCGCGACTACGTGAGCCCTTCTTCAAG 900
OY 898 ACCCTGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 957
DB 901 ACCCTGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
OY 958 GTGAGAAAGCCCAACCCCGACTGCAAGACCATCTGCGCGCTTCCGCGCGCGCGAGAG 1017
DB 961 GTGAGAAAGCCCAACCCCGACTGCAAGACCATCTGCGCGCTTCCGCGCGCGCGAGAG 1020
OY 1018 CTGAGAGATGATGACCGCTGCGAGGCGGTGGCGCGCGCGCGCGCGAGAGAGAGAG 1077
DB 1021 CTGAGAGATGATGACCGCTGCGAGGCGGTGGCGCGCGCGCGCGAGAGAGAGAGAG 1080
OY 1078 CTGCGCGAGCGATGAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1137
DB 1081 CTGCGCGAGCGATGAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1137
OY 1138 AAGGCGCGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1197
DB 1138 AAGGCGCGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1197
OY 1198 AACTGCGCGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1257
DB 1198 AACTGCGCGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1257
OY 1258 AAGGAGTGCACCGAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1317
DB 1258 AAGGAGTGCACCGAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1317
OY 1318 CGCCCGCGCACTTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1377
DB 1318 CGCCCGCGCACTTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1359
OY 1378 CCCCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1437
DB 1360 CCCCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1419
OY 1438 GACCGGAGCGCTACCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1497
DB 1420 G-----ACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1467
OY 1498 CTGAGCCAGTAA 1509
DB 1468 CTGAGCCAGTAA 1479

```

# RESULT 10

AAL44548  
ID AAL44548 standard; DNA; 1479 BP.

XX AAL44548;

XX AC  
XX 29-AUG-2003 (revised)  
DT 08-NOV-2002 (first entry)

XX DE  
XX HIV-1 p55gag polypeptide coding sequence 1.

XX KW HIV, ds; vaccine; gene; immune response; microparticle;  
KW adsorbent surface; poly(alpha-hydroxy acid); polyhydroxy butyric acid;  
KW polycaprolactone; polyorthoester; polycyanacrylate; detergent;  
KW submicron emulsion; viral infection; bacterial infection;  
KW parasitic infection; HIV-1 p55gag polypeptide.

XX OS Human immunodeficiency virus 1.

XX PN WO200226209-A2.

XX PD 04-APR-2002.

XX





AC ABL39953;  
XX  
DT 15-MAY-2002 (first entry)  
DE Synthetic Gag polynucleotide sequence SEQ ID NO:3.  
XX  
KW Human immunodeficiency virus type C; antigenic HIV type C protein;  
KW immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;  
KW immunostimulant; gene therapy; gene; ds.  
XX  
OS Human immunodeficiency virus; type C.  
OS Synthetic.  
XX  
PN WO200204493-A2.  
XX  
PD 17-JAN-2002.  
XX  
PF 05-JUL-2001; 2001WO-US021241.  
XX  
PR 05-JUL-2000; 2000US-00610313.  
XX  
PA (CHIR ) CHIRON CORP.  
PA (UYST-) UNIV STELLENBOSCH.  
XX  
PI Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg EJ,  
XX  
DR WPI; 2002-154920/20.  
XX  
PT New polynucleotides encoding antigenic HIV Type C polypeptides, useful in  
PT applications including DNA immunization or generation of packaging cell  
PT lines, particularly in gene therapy.  
XX  
PS Example 1; Fig 1; 233pp; English.  
XX  
CC The present invention describes expression cassettes comprising a  
CC polynucleotide sequence encoding a polypeptide comprising immunogenic HIV  
CC type C polypeptides. The expression cassettes comprise any of the HIV  
CC type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or Nef  
CC (1). (1) have immunostimulant activity and can be used in gene therapy.  
CC The HIV type C polynucleotides are useful in applications including DNA  
CC immunisation, generation of packaging cell lines, and production of HIV  
CC type C proteins. The polynucleotides are particularly useful in gene  
CC therapy and DNA immunisation applications. ABL39942 to ABL40054 and  
CC ABB06204 to ABB06215 represent sequences used in the exemplification of  
CC the present invention  
XX  
SQ Sequence 1479 BP; 325 A; 529 C; 463 G; 162 T; 0 U; 0 Other;  
Query Match 84.6%; Score 1276.8; DB 6; Length 1479;  
Best Local Similarity 92.5%; Pred. No. 1.2e-148;  
Matches 1399; Conservative 0; Mismatches 77; Indels 36; Gaps 4;  
QY 1 ATGGGCGCCCGCGCAGCATCTCTGCGCGGAGAGCTGCAAGTGGAGAGATCCGC 60  
DB 1 ATGGGCGCCCGCGCAGCATCTCTGCGCGGAGAGCTGCAAGTGGAGAGATCCGC 60  
QY 61 CTGCGCCCGCGCGGCAAGAGTGTACATGATGAAGCACTGTGTGGCCAGCCGCGAG 120  
DB 61 CTGCGCCCGCGCGGCAAGAGTGTACATGATGAAGCACTGTGTGGCCAGCCGCGAG 120  
QY 121 CTGAGGCGCTTCGCTGAAACCCCGCTGCTGAGACCGCGAGGCTGCAAGCAGATC 180  
DB 121 CTGAGGCGCTTCGCTGAAACCCCGCTGCTGAGACCGCGAGGCTGCAAGCAGATC 180  
QY 181 ATGAAGCAGCTGCAGCCGCTGAGACCGGACCGAGAGCTGCGAGCTGTACAC 240  
DB 181 ATCCGCGAGCTGCACCCGCTGAGACCGGACCGAGAGCTGGAAGCTGTTCAC 240  
QY 241 ACCGTGGCACCCTGTACTGTGTCAGCGCGCATGAGGTCCGCGACCAAGAGGCC 300  
DB 241 ACCGTGGCACCCTGTACTGTGTCAGCGAGAGATGAGGTCCGCGACCAAGAGGCC 300  
QY 301 CTGACAGATCGAGAGAGAGAGCAAGTCCAGCAGAGACCAGAGGCCAAGAG 360

DB 301 CTGACAGATCGAGAGAGAGAGCAAGTCCAGCAGAGATCCAGAGGCCAGGCC 360  
QY 361 GCCGAC---GGCAAGGTGAGCCAGACTACCCCATCTGTGACAGAACTTGAGGCCAGATG 417  
DB 361 GCCGACAGAGGCAAGGTGAGCCAGACTACCCCATCTGTGACAGAACTTGAGGCCAGATG 420  
QY 418 GTGCAACAGGCCATCAGCCCGGCAACCTGAACGCTGGTGAAGGTATCGAGAGAAG 477  
DB 421 GTGCAACAGGCCATCAGCCCGGCAACCTGAACGCTGGTGAAGGTATCGAGAGAAG 480  
QY 478 GCTTCAGCCCGAGGTGATCCCATGTTCAACCGCCCTGAGCGAGGGCGCACCCCCAG 537  
DB 481 GCTTCAGCCCGAGGTGATCCCATGTTCAACCGCCCTGAGCGAGGGCGCACCCCCAG 540  
QY 538 GACCTGAACAGATGTTGAACAACGTGGCGGCCACAGGCCGCTGAGATGCTGAAG 597  
DB 541 GACCTGAACAGATGTTGAACAACGTGGCGGCCACAGGCCGCTGAGATGCTGAAG 600  
QY 598 GACACCATCAACGAGAGAGCGCCGAGTGGGACCGCTGACACCCGCTGAGGCCGCC 657  
DB 601 GACACCATCAACGAGAGAGCGCCGAGTGGGACCGCTGACACCCGCTGAGGCCGCC 660  
QY 658 GTGGCCCCCGGCGAGATGCGGCAACCCCGCGCAGCAGCATGCGCGCACAGCAC 717  
DB 661 ATCGCCCCCGGCGAGATGCGGCAACCCCGCGCAGCAGCATGCGCGCACAGCAC 720  
QY 718 CTGACAGAGCAGATCGCTGATGACCAAGCAACCCCGCTGCGCGGCGCATCTAC 777  
DB 721 CTGACAGAGCAGATCGCTGATGACCAAGCAACCCCGCTGCGCGGCGCATCTAC 780  
QY 778 AAGCGGTGATCATCTGCGCTGACCAAGATGTGCGGATGTAACGCCCGTGAGCATC 837  
DB 781 AAGCGGTGATCATCTGCGCTGACCAAGATGTGCGGATGTAACGCCCGTGAGCATC 840  
QY 838 CTGACATCCCGCAGAGGCGCCCAAGAGAGCCCTTCCGCGACTACGTGAGACCGCTTCAAG 897  
DB 841 CTGACATCAAGCAGAGGCGCCCAAGAGAGCCCTTCCGCGACTACGTGAGACCGCTTCAAG 900  
QY 898 ACCCTGCGCGCGAGCAGAGGCCACCGAGACGTGAAGAACTGGATGACCGAGACCTGCTG 957  
DB 901 ACCCTGCGCGCGAGCAGAGGCCACCGAGAGGTGAAGAACTGGATGACCGAGACCTGCTG 960  
QY 958 GTGCAAGACGCCCAACCCCGACTGCAAGACCATCTGCGCGCTTCCGCGCGCGCCACC 1017  
DB 961 GTGCAAGACGCCCAACCCCGACTGCAAGACCATCTGCGCGCTTCCGCGCGCGCCACC 1020  
QY 1018 CTGAGAGAGATGATGACCGCTTCCAGAGGCGTGGGCGGCCCGCCACCAAGGCCCGCTG 1077  
DB 1021 CTGAGAGAGATGATGACCGCTTCCAGAGGCGTGGGCGGCCCGCCACCAAGGCCCGCTG 1080  
QY 1078 CTGGCGGAGCGATGAGCCAGGCGCAACAGCTGAACATCATGATGCAAGAGCAACTTC 1137  
DB 1081 CTGGCGGAGCGATGAGCCAGGCGCAACAGCTGAACATCATGATGCAAGAGCAACTTC 1137  
QY 1138 AAGGCGCCCGCGCGCAACGTCAAGTGTCTCAACTGCGCAAGAGGCGCCACTGCGCAAG 1197  
DB 1138 AAGGCGCCCGCGCGCATCTGTCAAGTGTCTCAACTGCGCAAGAGGCGCCACTGCGCGC 1197  
QY 1198 AACTGCGCGCGCGCGCAAGAGGCGTGTGAAGTGGCGCAAGAGGCGCCACAGATG 1257  
DB 1198 AACTGCGCGCGCGCGCGCAAGAGGCGTGTGAAGTGGCGCAAGAGGCGCCACAGATG 1257  
QY 1258 AAGACTGCACTGAGCGCGCAAGGCGCAACTTCTGCGCAAGATCTGGCCAGCCACAAGGCG 1317  
DB 1258 AAGACTGCACTGAGCGCGCAAGGCGCAACTTCTGCGCAAGATCTGGCCAGCCACAAGGCG 1317  
QY 1318 CGCCCGGCAACTTCTGCAAGAACCGCAGCGAGCCCGCGCCCAACCGTCCCAACCGCC 1377  
DB 1318 CGCCCGGCAACTTCTGCAAGAACCGCAGCGAGCCCGCGCCCAACCGTCCCAACCGCC 1377  
QY 1378 CCCCCCGGAGAGCTTCGCTTGAAGAGACCAACCCCGCCCAAGCAGAGGCCCAAG 1437  
DB 1378 CCCCCCGGAGAGCTTCGCTTGAAGAGACCAACCCCGCCCAAGCAGAGGCCCAAG 1437

Db 1360 CCCCCCGGAGAGCTTCCGCTTCGAGAGACCACCCCCGCCAGAGAGAGCAAG 1419  
QY 1438 GACCGGAGCCCTACCGGAGCCCTGACCGCCCTGCGACGCTGTTCGGACGGCCCC 1497  
Db 1420 G-----ACCGGAGACCTTGACCAAGCCTTGAGAGCCTGTTCGGCAAGACCCC 1467  
QY 1498 CTGAGCCAGTAA 1509  
Db 1468 CTGAGCCAGTAA 1479  
RESULT 12  
ADCL3239  
ID ADCL3239 standard; DNA; 3162 BP.  
XX AC ADCL3239;  
XX DT 18-DEC-2003 (first entry)  
XX DB DNA of HIV construct GagRTmut\_C SEQ ID NO 18.  
XX KW expression cassette; HIV Gag; Env; Int; Nef; p15RaseH; Pol; Tat; Prot;  
XX KW Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; ds.  
XX OS Human immunodeficiency virus.  
XX PN WO2003004620-A2.  
XX PD 16-JAN-2003.  
XX PF 05-JUL-2002; 2002WO-US021420.  
XX PR 05-JUL-2001; 2001US-0303192P.  
XX PR 31-AUG-2001; 2001US-0316860P.  
XX PR 16-JAN-2002; 2002US-0349871P.  
XX PA (CHIR ) CHIRON CORP.  
XX PA (UYST-) UNIV STBLEMBOSCH.  
XX PI Zur Megede J, Barnett SW, Lian Y, Engelbrecht S, Van Rensburg EJ;  
XX DR WPI; 2003-221593/21.  
XX PT New expression cassette comprising a polynucleotide sequence encoding a  
XX PT polypeptide including an HIV Gag, Env, Int, Nef, p15RaseH, Pol, Tat,  
XX PT Prot, or Rev polypeptide, useful for immunization, or generating  
XX PT packaging cell lines.  
XX PS Disclosure; Fig 15; 301pp; English.  
XX CC The invention relates to a novel expression cassette comprising a  
XX CC polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,  
XX CC Int, Nef, p15RaseH, Pol, Tat, Prot, or Rev polypeptide. The novel  
XX CC expression cassette can be used to treat HIV type C by gene therapy or  
XX CC used in the development of a vaccine. The gene delivery vector is  
XX CC administered intramuscularly, intramucosally, intranasally,  
XX CC subcutaneously, intradermally, transdermally, intravaginally,  
XX CC intrarectally, orally or intravenously. The expression cassette is useful  
XX CC for immunisation, generating packaging cell lines and producing HIV  
XX CC polypeptides. This polynucleotide sequence represents the DNA of an HIV  
XX CC Type C related sequence of the invention.  
SQ Sequence 3162 BP; 729 A; 1095 C; 975 G; 363 T; 0 U; 0 Other;  
Query Match 84.5%; Score 1275.2; DB 9; Length 3162;  
Best Local Similarity 92.5%; Pred. No. 1.6e-148;  
Matches 1388; Conservative 0; Mismatches 78; Indels 36; Gaps 4;

Db 67 CTGCGCCCCGGGCAAGAGTGTACATATGAAAGCACCTGTGTGGCCAGCCGCGAG 126  
QY 121 CTGAGGGCTTCCCTTGAACCCCGGCTCTGTGAGACCCGCGAGGGCTGCAAGATC 180  
Db 127 CTGAGAAATTCGCTTGAACCCCGGCTCTGTGAGACCCGCGAGGGCTGCAAGATC 186  
QY 181 ATGAAGACGCTGACGCCGCTGACAGCCGCGACCGAGAGCTGCGAGCTGTACAAC 240  
Db 187 ATCCGCAAGCTGACCCCGGCTGACAGCCGCGACCGAGAGCTGAAGAGCTGTCAAC 246  
QY 241 ACCGTGCAACCTGTACTGCGGTGACAGCCGCGATCGAGGTCCGACACCAAGAGGCC 300  
Db 247 ACCGTGCAACCTGTACTGCGGTGACAGAGATCGAGGTCCGACACCAAGAGGCC 306  
QY 301 CTGGAACAAGATCGAGAGAGCAACAAGTCCAGCAGAAAGCCAGAGCCAGAGAG 360  
Db 307 CTGGAACAAGATCGAGAGAGCAACAAGTCCAGCAGAAATCCAGAGGCCAGAGGCC 366  
QY 361 GCGGAC---GGCAAGTGAAGCCAGAACTACCCCATCTGCAAGACTGCAAGGCCAGATG 417  
Db 367 GCGGACAAAGGCAAGGTGAGCCAGAACTACCCCATCTGCAAGACTGCAAGGCCAGATG 426  
QY 418 GTGCAACCAAGCCATCAGCCCCCGCACCTGAACGCTGAGTGAAGTATCGAGAGAAG 477  
Db 427 GTGCAACCAAGCCATCAGCCCCCGCACCTGAACGCTGAGTGAAGTATCGAGAGAAG 486  
QY 478 GCCTTCAGCCCCGAGGTGATCCCATATGTTCAACGCTGAGCGAGGGCGCCACCCCCAG 537  
Db 487 GCCTTCAGCCCCGAGGTGATCCCATATGTTCAACGCTGAGCGAGGGCGCCACCCCCAG 546  
QY 538 GACCTGAACAAGATGTTGAACAACGCTGAGCGCGCAACAGGCCCATGCAAGTCTGAAG 597  
Db 547 GACCTGAACAAGATGTTGAACAACGCTGAGCGCGCAACAGGCCCATGCAAGTCTGAAG 606  
QY 598 GACACCATCAACAGAGAGGCGCGGAGTGGGAACGCTGCAACCCCTGCAAGGCCGCCCC 657  
Db 607 GACACCATCAACAGAGAGGCGCGGAGTGGGAACGCTGCAACCCCTGCAAGGCCGCCCC 666  
QY 658 GTGGCCCCCGGCGAGATGCGCGACCCCGCGGAGGACATCGCCGCGCCAGCAGCACC 717  
Db 667 ATCGCCCCCGGCGAGATGCGCGACCCCGCGGAGGACATCGCCGCGCCAGCAGCACC 726  
QY 718 CTGCAAGAGCAGATCGCTGTGATGACCAAGACCCCGCTGCGCGGACATCTAC 777  
Db 727 CTGCAAGAGCAGATCGCTGTGATGACCAAGACCCCGCTGCGCGGACATCTAC 786  
QY 778 AAGCGGTGATATCTCTGGGCTGAACAAGATGTGCGGATGTACAGCCCGCTGACATC 837  
Db 787 AAGCGGTGATATCTCTGGGCTGAACAAGATGTGCGGATGTACAGCCCGCTGACATC 846  
QY 838 CTGGAATATCCGCGAGGGGCCCAAGAGCCCTTTCGCGACTACGTGAGCCGCTTCAAG 897  
Db 847 CTGGAATATCAAGCAGGGGCCCAAGAGCCCTTTCGCGACTACGTGAGCCGCTTCAAG 906  
QY 898 ACCCTGCGCGCGAGCAGGCCACCAAGACGTGAAGAACTGATGACCGAACCTGCTG 957  
Db 907 ACCCTGCGCGCGAGCAGGCCACCAAGAGGTGAAGAACTGATGACCGAACCTGCTG 966  
QY 958 GTGCAAAACGCAACCCCGACTGCAAGAACCATCTGCGCGCTTTCGCGGCCCGGCCAAC 1017  
Db 967 GTGCAAAACGCAACCCCGACTGCAAGAACCATCTGCGCGCTTTCGCGGCCCGGCCAAC 1026  
QY 1018 CTGAGAGAGATGATGACCGCTGCGCAAGGCGGTGGCGGCCCGGCCCAAGGCCCGCTG 1077  
Db 1027 CTGAGAGAGATGATGACCGCTGCGCAAGGCGGTGGCGGCCCGGCCCAAGGCCCGCTG 1086  
QY 1078 CTGGCCGAGGCGATGAGCCAGGCCCAACAGCTGAACATCATGATGCAAGAGCAACTTC 1137  
Db 1087 CTGGCCGAGGCGATGAGCCAGGCCCAACAC--AGCTGATGATGCAAGAGCAACTTC 1143  
QY 1138 AAGGCCCCCCGCGCAACGTCAAGTCTTCAACTGCGGCAAGAGGGCCACATCGCCAAAG 1197

Db 1144 AAGGGCCCCCGGCATGTCAGTGTCTTCACTGCGGCAAGAGGGCCACATCGCCCGC 1203  
QY 1198 AACTGCGCGCCCCCGGCAAGAGGGCTGTGTAAGTGTGCGCAAGAGGGCCACAGATG 1257  
Db 1204 AACTGCGCGCCCCCGGCAAGAGGGCTGTGTAAGTGTGCGCAAGAGGGCCACAGATG 1263  
QY 1258 AAGGACTGACCGAGCGGCGCACTTCTGTGGGCAAGATTTGGCCCAAGCCCAAGGGC 1317  
Db 1264 AAGGACTGACCGAGCGGCGCACTTCTGTGGGCAAGATTTGGCCCAAGGGC 1323  
QY 1318 CGCCCGGCACTTCTGTGCAAGACCGGAGGCGCGCGCCCGCCACCGTGTCCCAACCGCC 1377  
Db 1324 CGCCCGGCACTTCTGTGAGAGCGG-----CGCCGAGCCCAACCGCC 1365  
QY 1378 CCCCCCGCGAGAGCTTCCGCTTCAAGAGACCAACCCCGCCCGCAAGAGAGAGCCCAAG 1437  
Db 1366 CCCCCCGCGAGAGCTTCCGCTTCAAGAGACCAACCCCGCGCAAGAGAGAGAGCAAG 1425  
QY 1438 GACCGCGAGCGCTTACCGCGAGCGCCCTGACCGCGCTTGGCGAGCGGCGCC 1497  
Db 1426 G-----ACCGCGAGAGCGCTGACCGCGCTGAAAGAGCGCTTGGCGAGAGCGCC 1473  
QY 1498 CTGAGCCAGTAA 1509  
Db 1474 CTGAGCCAGAAA 1485

## RESULT 13

ADCI3237 standard; DNA; 3462 BP.

ADCI3237  
ID ADCI3237 standard; DNA; 3462 BP.  
AC ADCI3237;  
XX 18-DEC-2003 (first entry)  
DT

DNA of HIV construct GagProteinRtmu\_C SEQ ID NO 16.

XX expression cassette; HIV Gag; Env; Int; Nef; p15RnaseH; Pol; Tat; Prot;  
KW Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; ds.

XX Human immunodeficiency virus.

OS WO2003004620-A2.

PN 16-JAN-2003.

PD 05-JUL-2002; 2002WO-US021420.

XX 05-JUL-2001; 2001US-0303192P.

PR 31-AUG-2001; 2001US-0316860P.

PR 16-JAN-2002; 2002US-0349871P.

XX (CHIR ) CHIRON CORP.

PA (UYST-) UNIV STELLENBOSCH.

XX Zur Megede J, Barnett SW, Lian Y, Engelbrecht S, Van Rensburg EJ;

PI WPI; 2003-221593/21.

XX New expression cassette comprising a polynucleotide sequence encoding a

PT polypeptide including an HIV Gag, Env, Int, Nef, p15RnaseH, Pol, Tat,

PT Prot, or Rev polypeptide, useful for immunization, or generating

XX packaging cell lines;

PS Disclosure; Fig 13; 301pp; English.

XX The invention relates to a novel expression cassette comprising a

CC polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,

CC Int, Nef, p15RnaseH, Pol, Tat, Prot, or Rev polypeptide. The novel

CC expression cassette can be used to treat HIV type C by gene therapy or

CC used in the development of a vaccine. The gene delivery vector is

CC administered intramuscularly, intranasally, intravenously,

CC subcutaneously, intradermally, transdermally, intravaginally,

CC intrarectally, orally or intravenously. The expression cassette is useful  
CC for immunisation, generating packaging cell lines and producing HIV  
CC polypeptides. This polynucleotide sequence represents the DNA of an HIV  
CC Type C related sequence of the invention.

SQ Sequence 3462 BP; 790 A; 1198 C; 1068 G; 406 T; 0 U; 0 Other;

Query Match 84.5%; Score 1275.2; DB 9; Length 3462;  
Best Local Similarity 92.5%; Pred. No. 1.6e-148;  
Matches 1398; Conservative 0; Mismatches 78; Indels 36; Gaps 4;

QY 1 ATGGGCGCCCGCGCAGATCTCTGCGCGGCAAGCTGCAAGTGGAGAGATCCGC 60.  
Db 7 ATGGGCGCCCGCGCAGATCTCTGCGCGGCAAGCTGCAAGTGGAGAGATCCGC 66  
QY 61 CTGCGCCCGCGCGCAAGAGCACTACATGTAAGCACTGTGTGGCCAGCGCGAG 120  
Db 67 CTGCGCCCGCGCGCAAGAGTGTCTACATATGAAGCACTGTGTGGCCAGCGCGAG 126  
QY 121 CTGAGGCTTGCCTGAAACCCCGGCTGTGAGACCGCGAGGGCTGCAAGCAGATC 180  
Db 127 CTGAGAAAGTTGCGCCCTGAACCCCGGCTGTGAGACCGCGAGGGCTGCAAGCAGATC 186  
QY 181 ATGAAGCACTGACGCCCGCTGACAGCGGCAACCGAGAGCTGGCAGCTGTACAAC 240  
Db 187 ATCCGCGAGCTGACCCCGCTGACAGACCGGCAAGAGAGCTGTGTTCAAC 246  
QY 241 ACCGTGCGCACCTGTACTGCTGTGACCGCGGCAATGAGGTCCGCGACCAAGAGGCC 300  
Db 247 ACCGTGCGCACCTGTACTGCTGTGACCGAGAAAGATGAGGTCCGCGACCAAGAGGCC 306  
QY 301 CTGACAAAGATCGAG 360  
Db 307 CTGACAAAGATCGAG 366  
QY 361 GCGGAC---GGCAAGGTGAGCCAGAACTTACCCCATGTGCAAGACCTGACAGGCGCAGATG 417  
Db 367 GCGGACAAAGGCAAGGTGAGCCAGAACTTACCCCATGTGCAAGACCTGACAGGCGCAGATG 426  
QY 418 GTGACACGAGGCTATCAGCCCGGCAACCTGAAAGCTGGTGAAGGTGATGAGAGAG 477  
Db 427 GTGACACGAGGCTATCAGCCCGGCAACCTGAAAGCTGGTGAAGGTGATGAGAGAG 486  
QY 478 GCCTTCAAGCCCGGAGGTATCCCATGTTCACCGGCTGAGCGAGGGCGGCAACCCCGCAG 537  
Db 487 GCCTTCAAGCCCGGAGGTATCCCATGTTCACCGGCTGAGCGAGGGCGGCAACCCCGCAG 546  
QY 538 GACCTGAACAGATGTTGAACACCGTGGCGGCGCAACGCGCGCATGAGATGCTGAAG 597  
Db 547 GACCTGAACAGATGTTGAACACCGTGGCGGCGCAACGCGCGCATGAGATGCTGAAG 606  
QY 598 GACACCATCAACGAGAGAGCGCGGAGTGGAGACCGCTGCAACCCCGTGAAGCGCGCGCC 657  
Db 607 GACACCATCAACGAGAGAGCGCGGAGTGGAGACCGCTGCAACCCCGTGAAGCGCGCGCC 666  
QY 658 GTGGCCCCCGGCGCAGATGCGGCGACCCCGGCGGAGAGATCGCGCGGCGCAACGAGACC 717  
Db 667 ATGCGCCCCCGGCGCAGATGCGGCGAGCCCGGCGGAGAGATCGCGCGGCGCAACGAGACC 726  
QY 718 CTGCAAGAGCAGATGCGCTGTGATGACCAACCCCGGCTGCGCGGCGGAGATCTAC 777  
Db 727 CTGCAAGAGCAGATGCGCTGTGATGACCAACCCCGGCTGCGCGGCGGAGATCTAC 786  
QY 778 AAGCGGTGATCATCTGCGGCTGAAACAAGATGTGCGGATGTACAGCCCGTGAAGCATC 837  
Db 787 AAGCGGTGATCATCTGCGGCTGAAACAAGATGTGCGGATGTACAGCCCGTGAAGCATC 846  
QY 838 CTGACATCCCGCAGAGGCGGCAAGAGCGCTTCCGCGACTAAGTGAAGCGCTTCTCAAG 897  
Db 847 CTGACATCAAGCAGAGGCGGCAAGAGCGCTTCCGCGACTAAGTGAAGCGCTTCTCAAG 906  
QY 898 ACCGTGCGCGCGAGAGAGCGGCAACCGAGAGCTGAAGACTGATGACCGAGACCTGTG 957



DB 907 ACCCTGCGCGCCGAGCAGAGCAACCAGAGGTGAAGAACTGATGACCGACACCTGCTG 966  
QY 958 GTGCAGAACGCCAACCCCGCATCTGCAAGACCATCTGCGCTCTCGGCCCCGGCGCACCC 1017  
XX |||||  
DB 967 GTGCAGAACGCCAACCCCGCATCTGCAAGACCATCTGCGCTCTCGGCCCCGGCGCACCC 1026  
QY 1018 CTGAGAGGAGATGATGACCGCTGCTGCAAGGCGCTGCGGCGGCGGCGGCGGCGGCGG 1077  
XX |||||  
DB 1027 CTGAGAGGAGATGATGACCGCTGCTGCAAGGCGCTGCGGCGGCGGCGGCGGCGG 1086  
QY 1078 CTGAGAGGAGATGATGACCGCTGCTGCAAGGCGCTGCTGCAAGGCGCTGCTGCAAG 1137  
DB 1087 CTGAGAGGAGATGATGACCGCTGCTGCAAGGCGCTGCTGCAAGGCGCTGCTGCAAG 1143  
QY 1138 AAGGCG 1197  
DB 1144 AAGGCG 1203  
QY 1198 AACTGCG 1257  
DB 1204 AACTGCG 1263  
QY 1258 AAGGAGCTGACCG 1317  
DB 1264 AAGGAGCTGACCG 1323  
QY 1318 CG 1377  
DB 1324 CG 1365  
QY 1378 CCG 1437  
DB 1366 CCG 1425  
QY 1438 GACCGCGAGCGCTACCG 1497  
DB 1426 G-----ACCGCGAGACCTGACGAGCGCTGAGAGAGCGCTGTTGCGCAAGCAACCC 1473  
QY 1498 CTGAGCCAGTAA 1509  
DB 1474 CTGAGCCAGAAA 1485

RESULT 14  
ACA03523 standard; DNA; 4419 BP.

AC ACA03523;

DT 22-MAY-2003 (first entry)

DB Synthetic DNA encoding immunogenic HIV peptide #6.

KW Immunogenic HIV polypeptide; human immunodeficiency virus; HIV; vaccine;  
KW gene therapy; packaging cell line; humoral immune response;  
KW cellular immune response; gene delivery vector; DNA immunisation; ds.

OS Synthetic.

PN WO2003004657-A1.

PD 16-JAN-2003.

PP 05-JUL-2002; 2002WO-US021421.

PR 05-JUL-2001; 2001US-0303192P.

PR 31-AUG-2001; 2001US-0316860P.

PR 16-JAN-2002; 2002US-0349728P.

PR 16-JAN-2002; 2002US-0349793P.

PR 16-JAN-2002; 2002US-0349871P.

PA (CHIR ) CHIRON CORP.

PI Zur Megede J, Barnett SW, Lian Y;  
XX WPI; 2003-221602/21.  
DR  
XX  
PT New synthetic polynucleotides encoding antigenic HIV type B and/or type C  
PT polypeptides, useful as immunogenic compositions or vaccines for  
PT generating humoral or cellular immune responses against HIV in a subject,  
PT especially humans.  
XX  
XX  
PS Example 1; Fig 11; 262pp; English.  
CC  
CC The invention describes a synthetic polynucleotide encoding 2 or more  
CC immunogenic HIV polypeptides, where at least 2 of the polypeptides are  
CC derived from different HIV subtypes. The polynucleotide is useful for  
CC immunisation, generation of packaging cell lines, or production of HIV  
CC polypeptides. The polynucleotide and its encoded proteins are useful as  
CC immunogenic compositions or vaccines for generating humoral or cellular  
CC immune responses against HIV in a subject, or for inducing neutralising  
CC antibodies against HIV. The gene delivery vector comprising the  
CC polynucleotide is also useful for DNA immunisation of, or for generating  
CC an immune response (e.g. a humoral or cellular immune response) in, a  
CC subject such as a mammal, particularly a human. This sequence encodes a  
CC human immunodeficiency virus immunogenic peptide  
XX  
SQ Sequence 4419 BP; 980 A; 1582 C; 1371 G; 486 T; 0 U; 0 Other;

Query Match 84.5%; Score 1275.2; DB 7; Length 4419;  
Best Local Similarity 92.5%; Pred. No. 1.5e-148;  
Matches 1398; Conservative 0; Mismatches 78; Indels 36; Gaps 4;

QY 1 ATGGGCG 60  
DB 7 ATGGGCG 66  
QY 61 CTGCG 120  
DB 67 CTGCG 126  
QY 121 CTGAGAGGCTTGGCG 180  
DB 127 CTGAGAGGCTTGGCG 186  
QY 181 ATGAAGCAGCTGACCG 240  
DB 187 ATCCGCGCAGCTGACCG 246  
QY 241 ACCGTGCGCAGCTGACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300  
DB 247 ACCGTGCGCAGCTGACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 306  
QY 301 CTGGAACAAGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360  
DB 307 CTGGAACAAGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 366  
QY 361 GCGGAC--GGCAAGGTGAGCGCAAGTACCCCATCTGTCAGAGGCGCGCGCGCGCG 417  
DB 367 GCGGACAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 426  
QY 418 GTGCAACAGGCGCATCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 477  
DB 427 GTGCAACAGGCGCATCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 486  
QY 478 GCCTTCAGCGCGCGAGGTGATCCCATGTTTCAACCGCGCGCGCGCGCGCGCGCGCGCG 537  
DB 487 GCCTTCAGCGCGCGAGGTGATCCCATGTTTCAACCGCGCGCGCGCGCGCGCGCGCGCG 546  
QY 538 GACTGAACAAGATGTTGAACAACCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 597  
DB 547 GACTGAACAAGATGTTGAACAACCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 606  
QY 598 GACACCATCAACGAGGAGGCGCGCGAGTGAACCGCTGCAACCGCGGTGACGGCGCGCC 657  
DB 607 GACACCATCAACGAGGAGGCGCGCGAGTGAACCGCTGCAACCGCGGTGACGGCGCGCC 666



```

QY 658 GTGCCCCCGCCAGATGCGGACCCCGCGCAGCGACATCGCGCGCCACGACACC 717
Db 667 ATGCCCCCGCCAGATGCGGAGCCCGCGCGCAGCATCGCGCGCACGACGACACC 726
QY 718 CTGCAAGAGAGATGCGCTGTGATGACCAACCCCGCTGCGCGCATCTAC 777
Db 727 CTGCAAGAGAGATGCGCTGTGATGACCAACCCCGCATCGCGCGCATCTAC 786
QY 778 AAGCGGTGATCATCTGCGCTGAAACAAGATCGTGAGATGACAGCCCGTGAAGCATC 837
Db 787 AAGCGGTGATCATCTGCGCTGAAACAAGATCGTGAGATGACAGCCCGTGAAGCATC 846
QY 838 CTGACATCGCGCCAGGCGCCCAAGAGAGCCCTTCGCGACTACGTGAGCCGCTTCTCAAG 897
Db 847 CTGACATCAAGAGAGGCGCCCAAGAGAGCCCTTCGCGACTACGTGAGCCGCTTCTCAAG 906
QY 898 ACCCTGCGCGCGCAGCAGGCGCCACGAGAGCTGAAAGACTGATGACCGAGACCCCTGCTG 957
Db 907 ACCCTGCGCGCGCAGCAGGCGCCACGAGAGCTGAAAGACTGATGACCGAGACCCCTGCTG 966
QY 958 GTGCAAGACCGCCAAACCCCGACTGCAAGACCATCTGCGCGCTCTGCGCGCGCCACACC 1017
Db 967 GTGCAAGACCGCCAAACCCCGACTGCAAGACCATCTGCGCGCTCTGCGCGCGCCACACC 1026
QY 1018 CTGAGAGAGATGATGACCGCTGCGCAGGCGCGCTGCGCGCGCGCCACAGGCGCGCGTG 1077
Db 1027 CTGAGAGAGATGATGACCGCTGCGCAGGCGCGCTGCGCGCGCGCCACAGGCGCGCGTG 1086
QY 1078 CTGCGCGAGCGGATGAGCGGCGCCACAGCGTGAACATCATGATGACAGAGACACTTC 1137
Db 1087 CTGCGCGAGCGGATGAGCGGCGCCACAGCGTGAACATCATGATGACAGAGACACTTC 1143
QY 1138 AAGGCGCGCGCGCGCAACGTCAAGTCTTCAACTGCGGCAAGAGAGGCGCCACTCGCCAG 1197
Db 1144 AAGGCGCGCGCGCGCATCGTCAAGTCTTCAACTGCGGCAAGAGAGGCGCCACTCGCCAG 1203
QY 1198 AACTGCGCGCGCGCGCGCAAGAGGCGTGTGAGAGTGCAGGCAAGAGGCGCCACAGATG 1257
Db 1204 AACTGCGCGCGCGCGCGCAAGAGGCGTGTGAGAGTGCAGGCAAGAGGCGCCACAGATG 1263
QY 1258 AAGGACTGACCGGAGCGCGCCCACTTCTGCGGCAAGATCTGCGCCACAGGCGCC 1317
Db 1264 AAGGACTGACCGGAGCGCGCCCACTTCTGCGGCAAGATCTGCGCCACAGGCGCC 1323
QY 1318 CGCGCGCGCACTTCTCTGCAAAACCGCAGCGCGCGCGCGCGCCACCGTGCACCGCGC 1377
Db 1324 CGCGCGCGCACTTCTCTGCAAAACCGCAGCGCGCGCGCGCGCGCCACCGTGCACCGCGC 1365
QY 1378 CCGCGCGCGGAGAGCTTCCGCTTTCAGAGAGACACCCCGCGCCCAAGCAGAGCGCCAG 1437
Db 1366 CCGCGCGCGGAGAGCTTCCGCTTTCAGAGAGACACCCCGCGCCCAAGCAGAGCGCCAG 1425
QY 1438 GACCGCGGAGCGCTTACCGCGAGCGCTTGAACCGCGCTTGCAGCGCTTGTGCGCAGCGCGCC 1497
Db 1426 G-----ACCGCGAGACCTTGAACCGCGCTTGAAGAGCGCTTGTGCGCAGCGCGCC 1473
QY 1498 CTGAGCGCAGTAA 1509
Db 1474 CTGAGCGCAGAAA 1485

```

```

RESULT 15
ADCl3240
ID ADCl3240 standard; DNA; 4419 BP.
XX AC ADCl3240;
XX DT 18-DEC-2003 (first entry)
XX DE DNA of HIV construct GagRTmutatRevNef_C SEQ ID NO 19.
XX expression cassette; HIV Gag; Env; Int; Nef; p15RnaseH; Pol; Tat; Prot;

```

```

KW Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; ds.
XX OS Human immunodeficiency virus.
XX PN WO2003004620-A2.
XX PD 16-JAN-2003.
XX PF 05-JUL-2002; 2002WO-US021420.
XX PR 05-JUL-2001; 2001US-0303192P.
XX PR 31-AUG-2001; 2001US-0316860P.
XX PR 16-JAN-2002; 2002US-0349871P.
XX PA (CHIR ) CHIRON CORP.
XX PA (UYST-) UNIV STELLENBOSCH.
XX PI Zur Megele J, Barnett SW, Lian Y, Engelbrecht S, Van Rensburg EJ;
XX DR MPL; 2003-221593/21.
XX PT New expression cassette comprising a polynucleotide sequence encoding a
PT polypeptide including an HIV Gag, Env, Int, Nef, p15RnaseH, Pol, Tat,
PT Prot, or Rev polypeptide, useful for immunization, or generating
PT packaging cell lines.
XX PS Disclosure; Fig 16; 301pp; English.
XX CC The invention relates to a novel expression cassette comprising a
CC polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,
CC Int, Nef, p15RnaseH, Pol, Tat, Prot, or Rev polypeptide. The novel
CC expression cassette can be used to treat HIV type C by gene therapy or
CC used in the development of a vaccine. The gene delivery vector is
CC administered intramuscularly, intramuscularly, intravenously,
CC subcutaneously, intradermally, transdermally, intravaginally,
CC intrarectally, orally or intravenously. The expression cassette is useful
CC for immunisation, generating packaging cell lines and producing HIV
CC polypeptides. This polynucleotide sequence represents the DNA of an HIV
CC Type C related sequence of the invention.
SQ Sequence 4419 BP; 980 A; 1582 C; 1371 G; 486 T; 0 U; 0 Other;
Query Match 84.5%; Score 1275.2; DB 9; Length 4419;
Best Local Similarity 92.5%; Pred. No. 1.5e-148;
Matches 1398; Conservative 0; Mismatches 78; Indels 36; Gaps 4;
QY 1 ATGGCGCGCGCGCCAGCATCTGCGCGCGCGGAGAGCTGGAACAAGTGGAGAAATCCGC 60
Db 7 ATGGCGCGCGCGCGCCAGCATCTGCGCGCGCGGAGAGCTGGAACAAGTGGAGAAATCCGC 66
QY 61 CTGCGCGCGCGCGCGCAAGAGCACTACATGCTGAAGCACTGTGTGCGCGCAGCGCGAG 120
Db 67 CTGCGCGCGCGCGCGCAAGAGCTACTATGATGAAGCACTGTGTGCGCGCAGCGCGAG 126
QY 121 CTGAGAGGCTTTCGCTTGAACCCCGCGCTGTGAGAGACCGCGGAGGCTGCAAGCATC 180
Db 127 CTGAGAGAGTTCGCTTGAACCCCGCGCTGTGAGAGACCGCGGAGGCTGCAAGCATC 186
QY 181 ATGAAGCAGCTGCAAGCGCGCGCTGCAAGCGCGGAGAGGCTGCGCAGCGCTGTACAAC 240
Db 187 ATCCGCCAGCTGCAAGCGCGCGCTGCAAGCGCGGAGAGGCTGTGAAGAGCTGTTCAC 246
QY 241 ACCGTGCGCACCTGTACTGTGCGTCAAGCGCGGATGAGGTCGCGGACACCAAGAGGCC 300
Db 247 ACCGTGCGCACCTGTACTGTGCGTCAAGCGGATGAGGTCGCGGACACCAAGAGGCC 306
QY 301 CTGACAAGATGAGAGAGAGAGAGAGCAAGTCCAGAGAGAGAGAGAGAGAGAGAGAG 360
Db 307 CTGACAAGATGAGAGAGAGAGAGAGAGCAAGTCCAGAGAGATTCAGAGAGAGAGAGCC 366
QY 361 GCCGAC---GCCAAGGTGAGCGCAAGTACCCTATCGTGCAGAACTTGAGGCGCAGATG 417
Db 367 GCCGACAAGGCGCAAGGTGAGCGCAAGTACCCTATCGTGCAGAACTTGAGGCGCAGATG 426

```

QY 418 GTGCACCAAGGCGATCAGCCCCCGCAACCTTGAACGCTGAGTGAAGTGAATCGAGAGAAG 477  
DB 427 GTGCACCAAGGCGATCAGCCCCCGCAACCTTGAACGCTGAGTGAAGTGAATCGAGAGAAG 486  
QY 478 GCCTTCAAGCGCGAGAGTGAATCCCATATGTTCAACCGCCCTGAGCGAGGCGCGCAACCCCGCAG 537  
DB 487 GCCTTCAAGCGCGAGAGTGAATCCCATATGTTCAACCGCCCTGAGCGAGGCGCGCAACCCCGCAG 546  
QY 538 GACCTGAACACGATGTTGAACACCTGAGCGCGCAACGAGCGCGCATGAGATGCTGAAG 597  
DB 547 GACCTGAACACGATGTTGAACACCTGAGCGCGCAACGAGCGCGCATGAGATGCTGAAG 606  
QY 598 GACACCATCAACGAGAGGCGCGGAGTGGACCGCGCTGCAACCGCTGAGCGCGCGCC 657  
DB 607 GACACCATCAACGAGAGGCGCGGAGTGGACCGCGCTGCAACCGCTGAGCGCGCGCC 666  
QY 658 GTGGCCCCCGCGCAGATGCGCGCAACCCCGCGCGAGCATGCGCGCGCGCGCAACGAGCAAC 717  
DB 667 ATGCGCCCCCGCGCAGATGCGCGCAACCCCGCGCGAGCATGCGCGCGCGCGCAACGAGCAAC 726  
QY 718 CTGCAGAGAGCAGATCGCTTGAATGACAGCAACCCCGCGCGAGCTGCGCGCGCGCATCTAC 777  
DB 727 CTGCAGAGAGCAGATCGCTTGAATGACAGCAACCCCGCGCGAGCTGCGCGCGCGCATCTAC 786  
QY 778 AAGCGGTGATCATCTGAGCGCTGAAACAGATGCTGCGAGTGAACAGCCCGTGAAGCATC 837  
DB 787 AAGCGGTGATCATCTGAGCGCTGAAACAGATGCTGCGAGTGAACAGCCCGTGAAGCATC 846  
QY 838 CTGACATCCCGCAGAGCG 897  
DB 847 CTGACATCCCGCAGAGCG 906  
QY 898 ACCCTGCG 957  
DB 907 ACCCTGCG 966  
QY 958 GTGCAGAAACGCAACCCCGCACTGCAAGACCATCTGCGCGCTCTGCGCGCGCGCGCGCGCGCG 1017  
DB 967 GTGCAGAAACGCAACCCCGCACTGCAAGACCATCTGCGCGCTCTGCGCGCGCGCGCGCGCGCG 1026  
QY 1018 CTGAGAGAGATGATGACCGCTGCG 1077  
DB 1027 CTGAGAGAGATGATGACCGCTGCG 1086  
QY 1078 CTGCGCGAGCGATGAGCG 1137  
DB 1087 CTGCGCGAGCGATGAGCG 1143  
QY 1138 AAGGGCG 1197  
DB 1144 AAGGGCG 1203  
QY 1198 AACTGCG 1257  
DB 1204 AACTGCG 1263  
QY 1258 AAGGACTGACCGAGCG 1317  
DB 1264 AAGGACTGACCGAGCG 1323  
QY 1318 CG 1377  
DB 1324 CG 1385  
QY 1378 CCCCCCGCGAGAGCTTCGCTTGAAGAGACCAACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1437  
DB 1386 CCCCCCGCGAGAGCTTCGCTTGAAGAGACCAACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1445  
QY 1438 GACCGCGAGCG 1497  
DB 1446 G-----ACCGGAGAGCTTGAACGAGCTTGAAGAGCTTGAAGAGCTTGAAGAGCTTGAAGAGCTT 1503

QY 1498 CTGAGCCAGTAA 1509  
DB 1474 CTGAGCCAGTAA 1485

Search completed: June 1, 2004, 11:31:51  
Job time : 621 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 1, 2004, 10:26:25 ; Search time 5848 Seconds  
(without alignments)  
11184.104 Million cell updates/sec

Title: US-09-475-704A-4

Perfect score: 1509

Sequence: 1 atggggcgccgcgcagcat.....gcggccctgagccagtaa 1509

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBankl:\*

1: gb\_ba:\*

2: gb\_htg:\*

3: gb\_in:\*

4: gb\_om:\*

5: gb\_ov:\*

6: gb\_pat:\*

7: gb\_ph:\*

8: gb\_pl:\*

9: gb\_pr:\*

10: gb\_ro:\*

11: gb\_sts:\*

12: gb\_sy:\*

13: gb\_un:\*

14: gb\_vi:\*

15: em\_ba:\*

16: em\_fun:\*

17: em\_hum:\*

18: em\_in:\*

19: em\_mu:\*

20: em\_om:\*

21: em\_or:\*

22: em\_ov:\*

23: em\_pat:\*

24: em\_ph:\*

25: em\_pl:\*

26: em\_ro:\*

27: em\_sts:\*

28: em\_un:\*

29: em\_vi:\*

30: em\_htg\_hum:\*

31: em\_htg\_inv:\*

32: em\_htg\_other:\*

33: em\_htg\_mus:\*

34: em\_htg\_pln:\*

35: em\_htg\_rod:\*

36: em\_htg\_mam:\*

37: em\_htg\_vrt:\*

38: em\_sy:\*

39: em\_htgo\_hum:\*

40: em\_htgo\_mus:\*

41: em\_htgo\_other:\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1509	100.0	1509	6	AX455888	AX455888 Sequence
2	1507.4	99.9	1509	6	AX468544	AX468544 Sequence
3	1504.2	99.7	1509	6	AX468548	AX468548 Sequence
4	1489.8	98.7	1509	6	AX455905	AX455905 Sequence
5	1297.6	86.0	1494	6	AX455935	AX455935 Sequence
6	1294.8	85.8	1491	6	AX455983	AX455983 Sequence
7	1291.4	85.6	1485	12	AY181195	AY181195 Synthetic
8	1276.8	84.6	1479	6	AX455887	AX455887 Sequence
9	1276.8	84.6	1479	6	AX468543	AX468543 Sequence
10	1271.2	84.2	1479	6	AX468547	AX468547 Sequence
11	1260.8	83.6	1479	6	AX455904	AX455904 Sequence
12	1202.2	79.7	4288	6	AX149648	AX149648 Sequence
13	1201	79.6	1509	12	AF201927	AF201927 Synthetic
14	1201	79.6	1515	6	BD263637	BD263637 Improved
15	1201	79.6	1515	6	AR373320	AR373320 Sequence
16	1201	79.6	4472	6	BD263697	BD263697 Improved
17	1201	79.6	4472	6	AR373380	AR373380 Sequence
18	1201	79.6	4608	6	BD263698	BD263698 Improved
19	1201	79.6	4608	6	AR373381	AR373381 Sequence
20	1201	79.6	4689	6	BD263696	BD263696 Improved
21	1201	79.6	4689	6	AR373379	AR373379 Sequence
22	1201	79.6	4766	6	BD263695	BD263695 Improved
23	1201	79.6	4766	6	AR373378	AR373378 Sequence
24	1199	79.5	2031	6	BD263640	BD263640 Improved
25	1199	79.5	2031	6	AR373323	AR373323 Sequence
26	1186	78.6	9166	6	AX427930	AX427930 Sequence
27	1170.8	77.6	9788	6	AX427936	AX427936 Sequence
28	1145	75.9	1539	6	AX306431	AX306431 Sequence
29	1143.2	75.8	1548	12	AF287354	AF287354 Synthetic
30	1119	74.2	6438	6	AX427919	AX427919 Sequence
31	1119	74.2	9194	6	AX427925	AX427925 Sequence
32	1119	74.2	9194	6	AX427926	AX427926 Sequence
33	1119	74.2	12411	6	AX427927	AX427927 Sequence
34	1117.8	74.1	1847	12	AF202464	AF202464 Synthetic
35	1117.8	74.1	1847	12	AF202465	AF202465 Synthetic
36	1117.8	74.1	1853	6	BD263638	BD263638 Improved
37	1117.8	74.1	1853	6	AR373321	AR373321 Sequence
38	1117.8	74.1	1865	6	BD263700	BD263700 Improved
39	1117.8	74.1	1865	6	BD263701	BD263701 Improved
40	1117.8	74.1	1865	6	AR373383	AR373383 Sequence
41	1117.8	74.1	1865	6	AR373384	AR373384 Sequence
42	1117.8	74.1	4319	6	BD263639	BD263639 Improved
43	1117.8	74.1	4319	6	AR373322	AR373322 Sequence
44	1103.8	73.1	9170	6	AX427928	AX427928 Sequence
45	1103.8	73.1	9189	6	AX427921	AX427921 Sequence

#### ALIGNMENTS

RESULT 1  
AX455888  
LOCUS AX455888 1509 bp DNA linear PAT 06-JUL-2002  
DEFINITION Sequence 4 from Patent WO0204493.  
ACCESSION AX455888  
VERSION AX455888.1 GI:21714881  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1 zur Mege, J., Barnett, S.W., Engelbrecht, S. and van Rensburg, B.  
AUTHORS Polynucleotides encoding antigenic hiv type c polypeptides,  
TITLE Polypeptides and uses thereof  
JOURNAL Patent: WO 0204493-A 4 17-JAN-2002;

Pred. No. is the number of results predicted by chance to have a

```

CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)
location/Qualifiers
1.1509
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="synthetic Gag of HIV strain AF110967"

```

**ORIGIN**

Query Match	100.0%;	Score 1509;	DB 6;	Length 1509;
Best Local Similarity	100.0%;	Pred. No. 1.3e-150;		
Matches 1509; Conservative	0;	Mismatches	0;	Gaps 0;

OY	1	ATGGGCGCCCGCCAGCATCTCTGGCGGGCGAAGACTGGACAAATGGGAGAAAGATCCG	60
Db	1	ATGGGCGCCCGCCAGCATCTCTGGCGGGCGAAGACTGGACAAATGGGAGAAAGATCCG	60
OY	61	CTGCGCCCCCGCGCAAGAGCACTACATGCTGAAGCACTGTGTGGGCCAGCCCGAG	120
Db	61	CTGCGCCCCCGCGCAAGAGCACTACATGCTGAAGCACTGTGTGGGCCAGCCCGAG	120
OY	121	CTGGAAGGCTTCCGCTTGAAACCCCGGCTGTGTGAGACCGCCGAGGCTGCAAGCAGATC	180
Db	121	CTGGAAGGCTTCCGCTTGAAACCCCGGCTGTGTGAGACCGCCGAGGCTGCAAGCAGATC	180
OY	181	ATGAAGCAGCTGCAACCCCGCTGCAAGCCCGCACCGAGAGCTGGCAGCCTGTACAAC	240
Db	181	ATGAAGCAGCTGCAACCCCGCTGCAAGCCCGCACCGAGAGCTGGCAGCCTGTACAAC	240
OY	241	ACCGTGGCCACCTGTACTGTGCTGCAACCGCGGCATCGAGGTCCGCGACACCAAGAGGCC	300
Db	241	ACCGTGGCCACCTGTACTGTGCTGCAACCGCGGCATCGAGGTCCGCGACACCAAGAGGCC	300
OY	301	CTGACAAGATCGAAGAGGAGCAGAACAAAGTCCCAAGACAGAACCCAGGCCAAGAG	360
Db	301	CTGACAAGATCGAAGAGGAGCAGAACAAAGTCCCAAGACAGAACCCAGGCCAAGAG	360
OY	361	GCCGACGCGCAAGGTGAGCCAGAACTACCCCATCTGTGCAGAACCTGACAGGGCCAGATGTTG	420
Db	361	GCCGACGCGCAAGGTGAGCCAGAACTACCCCATCTGTGCAGAACCTGACAGGGCCAGATGTTG	420
OY	421	CACCAGGCCATCAGCCCCCGCAACCTGAACGCTGTGGTGAAGGTATCGAGGAGAAAGCC	480
Db	421	CACCAGGCCATCAGCCCCCGCAACCTGAACGCTGTGGTGAAGGTATCGAGGAGAAAGCC	480
OY	481	TTCAAGCCCGAGGTGATCCCATGTTCAACCGCTGTGAGCGAGGGCCACCCCGCAGAC	540
Db	481	TTCAAGCCCGAGGTGATCCCATGTTCAACCGCTGTGAGCGAGGGCCACCCCGCAGAC	540
OY	541	CTGAACACGATTTGAACACCGTGGGCGGCCACCAAGGCCGCATGACATGCTGAAGAC	600
Db	541	CTGAACACGATTTGAACACCGTGGGCGGCCACCAAGGCCGCATGACATGCTGAAGAC	600
OY	601	ACCATCAACGAGAGGCGCGCAGTGGGACCGCTGCAACCCGCTGACGCGCCCGGTG	660
Db	601	ACCATCAACGAGAGGCGCGCAGTGGGACCGCTGCAACCCGCTGACGCGCCCGGTG	660
OY	661	GCCCCCGGCCAGATGCGCGACCCCGCGCGCAGCGACATCGCCGGCGCAACAGCACCTTG	720
Db	661	GCCCCCGGCCAGATGCGCGACCCCGCGCGCAGCGACATCGCCGGCGCAACAGCACCTTG	720
OY	721	CAGGAGCAGATCGCTTGATGACCAAGCAACCCCGCTGCGCCGTGGCGCAGATCTACAAG	780
Db	721	CAGGAGCAGATCGCTTGATGACCAAGCAACCCCGCTGCGCCGTGGCGCAGATCTACAAG	780
OY	781	CGGTGATCATCTTGGGCTTGAACAAGATCGTGCAGATGTACAAGCCCTGTAGCATCTTG	840
Db	781	CGGTGATCATCTTGGGCTTGAACAAGATCGTGCAGATGTACAAGCCCTGTAGCATCTTG	840
OY	841	GACATCCGCCAGGGCCCCCAAGGAGCTTCCGCGACTACGTGACCGCTTCTTCAAGACC	900
Db	841	GACATCCGCCAGGGCCCCCAAGGAGCTTCCGCGACTACGTGACCGCTTCTTCAAGACC	900

QY	901	CTGCGCGCCGAGCAGGCGCA	CC	CAGGACGTGAAGAACTG	ATGA	TGACCGGAGACCC	TGCTGTG	960
DB	901	CTGCGCGCCGAGCAGGCGCA	CC	CAGGACGTGAAGAACTG	ATGA	TGACCGGAGACCC	TGCTGTG	960
QY	961	CAGAACGCCAACCCCGACTG	CA	AGACCATCTTGCGCGCTCT	TGGCCCGCCGCGCCAC	CCCTG	1020	
DB	961	CAGAACGCCAACCCCGACTG	CA	AGACCATCTTGCGCGCTCT	TGGCCCGCCGCGCCAC	CCCTG	1020	
QY	1021	GAGGAGATGATGACCGCCTT	CC	ACGAGGGCGCGGCCCGCCG	CCCAAGCCCGCGTCTG		1080	
DB	1021	GAGGAGATGATGACCGCCTT	CC	ACGAGGGCGCGGCCCGCCG	CCCAAGCCCGCGTCTG		1080	
QY	1081	GCCGAGGCGCATGAGCCAG	CGCCAA	CAGCGCTGAACATCAT	TGATGCA	GAAAGCAACTTCA	1140	
DB	1081	GCCGAGGCGCATGAGCCAG	CGCCAA	CAGCGCTGAACATCAT	TGATGCA	GAAAGCAACTTCA	1140	
QY	1141	GCCCCCCCCGCGCAACGT	CA	AGTGCTTCAACTGCGGCA	AGAGGGCCACATCG	CCAAAGAAC	1200	
DB	1141	GCCCCCCCCGCGCAACGT	CA	AGTGCTTCAACTGCGGCA	AGAGGGCCACATCG	CCAAAGAAC	1200	
QY	1201	TGCGCGCGCCCCCGCA	AGAGGGCTGCTG	GAAGTGC	CGCAAGAGGGCCAC	CAGATGAAG	1260	
DB	1201	TGCGCGCGCCCCCGCA	AGAGGGCTGCTG	GAAGTGC	CGCAAGAGGGCCAC	CAGATGAAG	1260	
QY	1261	GACTGCACCGAGCGCCAG	CGCAACTTCTCTG	GGCAAGATCTG	GCCACGACCA	CAAGGCGCG	1320	
DB	1261	GACTGCACCGAGCGCCAG	CGCAACTTCTCTG	GGCAAGATCTG	GCCACGACCA	CAAGGCGCG	1320	
QY	1321	CCCGGCAACTTCTTG	CAGAACCG	CAGCGAGCCCGCGCGCC	CCACCGTGCCAC	CGCCCCC	1380	
DB	1321	CCCGGCAACTTCTTG	CAGAACCG	CAGCGAGCCCGCGCGCC	CCACCGTGCCAC	CGCCCCC	1380	
QY	1381	CCCGCGGAGAGCTTCCG	CTTGCAGAGACCA	CCCCCGCCCCCA	AGCAGAGCCCA	AGGAC	1440	
DB	1381	CCCGCGGAGAGCTTCCG	CTTGCAGAGACCA	CCCCCGCCCCCA	AGCAGAGCCCA	AGGAC	1440	
QY	1441	CGCGAGCCCTAACCG	CGAGCCCC	TGACCGCCCTG	CGCAGCCTGTT	CGGCAAGGGCCCCCTG	1500	
DB	1441	CGCGAGCCCTAACCG	CGAGCCCC	TGACCGCCCTG	CGCAGCCTGTT	CGGCAAGGGCCCCCTG	1500	
QY	1501	AGCCAGTAA	1509					
DB	1501	AGCCAGTAA	1509					

RESULT 2  
AX468544

DEFINITION	Sequence 64 from Patent WO0226209.
ACCESSION	AY468544
VERSION	AY468544.1 GI:21901374

**SOURCE** Human immunodeficiency virus 1 (HIV-1)

**ORGANISM** Human immunodeficiency virus 1  
Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate  
lentivirus group.

REFERENCE	AUTHORS	TITLE	JOURNAL
1	O'Hagan, D., Otten, G., Donnelly, J.J., Polo, J.M., Barnett, S., Singh, M., Ulmer, J. and Dubelsky, T.W.	Microparticles for delivery of the heterologous nucleic acids	Patent: WO 0226209-A 64 04-APR-2002;

FEATURES	Location/Qualifiers
SOURCE	1. .1509

```

      .
      /organism="Human immunodeficiency virus 1"
      /mol_type="unassigned DNA"
      /db_xref="taxon:11676"

```

Query Match	99.9%	Score 1507.4;	DB 6;	Length 1509;
Best Local Similarity	99.9%	Pred. No. 1.9e-150;		
Matches 1508; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

```

QY 1 ATGGGCGCGCGCGCAGCATCTCTGCGCGCGGAGAGCTGSAACAAGTGGAGATCCGC 60
DB 1 ATGGGCGCGCGCGCAGCATCTCTGCGCGCGGAGAGCTGSAACAAGTGGAGATCCGC 60
QY 61 CTGCGCGCGCGCGCAGAGAGCACTACATGCTGAAGCACTGTGTGTGGCCAGCCGCGAG 120
DB 61 CTGCGCGCGCGCGCAGAGAGCACTACATGCTGAAGCACTGTGTGTGGCCAGCCGCGAG 120
QY 121 CTGAGAGGCTTCTGCTGTAACCTCCGCGCTGTGAGAGACCGCGGAGGGCTGCAAGCATC 180
DB 121 CTGAGAGGCTTCTGCTGTAACCTCCGCGCTGTGAGAGACCGCGGAGGGCTGCAAGCATC 180
QY 181 ATGAAGCAGCTGACAGCCCGCTGCAAGCCGCGACCGAGAGCTGCGCAGCTGTACAAC 240
DB 181 ATGAAGCAGCTGACAGCCCGCTGCAAGCCGCGACCGAGAGCTGCGCAGCTGTACAAC 240
QY 241 ACCGTGGCACTCTGTACTGCTGTCAGACCGCGCATCGAGGTCCGACCAAGAGGCC 300
DB 241 ACCGTGGCACTCTGTACTGCTGTCAGACCGCGCATCGAGGTCCGACCAAGAGGCC 300
QY 301 CTGACAAAGATCGAGAGAGAGAGCAACAGTCCAGCAGAAAGCCAGCAGGCCAAGAG 360
DB 301 CTGACAAAGATCGAGAGAGAGAGCAACAGTCCAGCAGAAAGCCAGCAGGCCAAGAG 360
QY 361 GCGGACGCAAGGTGAGCCAGAACTACCCCATCTGTGAGAACTGTGAGGGCCAGATGCTG 420
DB 361 GCGGACGCAAGGTGAGCCAGAACTACCCCATCTGTGAGAACTGTGAGGGCCAGATGCTG 420
QY 421 CACCAGGCCATCAGCCCCCGCACCTGTAAACGCTGGGTGAAGGTGATCGAGAGAAAGGCC 480
DB 421 CACCAGGCCATCAGCCCCCGCACCTGTAAACGCTGGGTGAAGGTGATCGAGAGAAAGGCC 480
QY 481 TTCAGCCCCGAGGTGATCCCATGTTCAACCCCTGAGCGAGGGCGCACCCCGCAGAG 540
DB 481 TTCAGCCCCGAGGTGATCCCATGTTCAACCCCTGAGCGAGGGCGCACCCCGCAGAG 540
QY 541 CTGAACACGATGTTGAACAACGCTGGGCGCGCACAGGCCCGCATGCAATGCTGAAGGAC 600
DB 541 CTGAACACGATGTTGAACAACGCTGGGCGCGCACAGGCCCGCATGCAATGCTGAAGGAC 600
QY 601 ACCATCAACGAGAGAGCGCGCGAGTGGAGCGCTGCAACCCCGTGCAGGCCGCCCGTG 660
DB 601 ACCATCAACGAGAGAGCGCGCGAGTGGAGCGCTGCAACCCCGTGCAGGCCGCCCGTG 660
QY 661 GCGCGCGCGCAGATGCGCGACCCCGCGGAGCAGCATCGCGCGCGCAGCAGCACCCCTG 720
DB 661 GCGCGCGCGCAGATGCGCGACCCCGCGGAGCAGCATCGCGCGCGCAGCAGCACCCCTG 720
QY 721 CAGGAGCAGATCGCTGATGACAGCAACCCCGCTGCGCGGTGGCGCATCTACAAG 780
DB 721 CAGGAGCAGATCGCTGATGACAGCAACCCCGCTGCGCGGTGGCGCATCTACAAG 780
QY 781 CGGTGATCATCTGCGCTGAAACAAGATCGTGGATGTACAGCCCGTGAAGCATCTG 840
DB 781 CGGTGATCATCTGCGCTGAAACAAGATCGTGGATGTACAGCCCGTGAAGCATCTG 840
QY 841 GACATCCGCGCAGGCGCGCGAGAGCGCTTCGCGACTACGTGAGCCGCTTCTCAAGACC 900
DB 841 GACATCCGCGCAGGCGCGCGAGAGCGCTTCGCGACTACGTGAGCCGCTTCTCAAGACC 900
QY 901 CTGCGCGCGCAGAGAGAGCCAGAGACGTAAGAACTGATGACCGAGACCTGTGCTG 960
DB 901 CTGCGCGCGCAGAGAGAGCCAGAGACGTAAGAACTGATGACCGAGACCTGTGCTG 960
QY 961 CAGAGCGCAACCCCGACTGCAAGACACTCTGCGCGCTCTCGCGCCCGCGGCGCACCCCTG 1020
DB 961 CAGAGCGCAACCCCGACTGCAAGACACTCTGCGCGCTCTCGCGCCCGCGGCGCACCCCTG 1020
QY 1021 GAGGAGATGATGACCGCTGCGCAGGGCGTGGGGCGCGCCAGAGGCCCGCGTGTCTG 1080
DB 1021 GAGGAGATGATGACCGCTGCGCAGGGCGTGGGGCGCGCCAGAGGCCCGCGTGTCTG 1080
QY 1081 GCGGAGCGGATGAGCGAGGCCCAACAGCGTGAACATCATGATGCAAGAGCAACTTCAAG 1140

```

```

DB 1081 GCGGAGCGGATGAGCCAGGCCAACAGCGTGAACATCATGATGCAAGAGAGCAACTTCAAG 1140
QY 1141 GCGCGCGCGCGCAACGTCAGAGCTTCAACTGCGCGCAGAGAGGCCACATCGCAGAAAC 1200
DB 1141 GCGCGCGCGCGCAACGTCAGAGCTTCAACTGCGCGCAGAGAGGCCACATCGCAGAAAC 1200
QY 1201 TGCGCGCGCGCGCGCAGAGAGGCTGTGAGAGTGCAGAGAGGGCCACCAAGTGAAG 1260
DB 1201 TGCGCGCGCGCGCGCAGAGAGGCTGTGAGAGTGCAGAGAGGGCCACCAAGTGAAG 1260
QY 1261 GACTGCAACGAGCGCCAGAGCACTTCTGCGCAAGATCTGCGCCAGCCACAGAGGCCGC 1320
DB 1261 GACTGCAACGAGCGCCAGAGCACTTCTGCGCAAGATCTGCGCCAGCCACAGAGGCCGC 1320
QY 1321 CCGCGCACTTCTTGCAGAACCGCAGCGAGCCCGCGCCCGCACCGTCCACCGCCCC 1380
DB 1321 CCGCGCACTTCTTGCAGAACCGCAGCGAGCCCGCGCCCGCACCGTCCACCGCCCC 1380
QY 1381 CCGCGCGAGAGCTTCCGCTTGCAGAGAGACCAACCCCGCGCGCGCAGAGGCCAAGGAC 1440
DB 1381 CCGCGCGAGAGCTTCCGCTTGCAGAGAGACCAACCCCGCGCGCGCAGAGGCCAAGGAC 1440
QY 1441 CCGGAGCCCTACCGCGAGCCCTGACCGCCCTGCGGAGCCTGTTGCGCAGCGGCCCTG 1500
DB 1441 CCGGAGCCCTACCGCGAGCCCTGACCGCCCTGCGGAGCCTGTTGCGCAGCGGCCCTG 1500
QY 1501 AGCCAGTAA 1509
DB 1501 AGCCAGTAA 1509

RESULT 3
AX468548 1509 bp DNA linear PAT 16-JUL-2002
LOCUS AX468548
DEFINITION Sequence 68 from Patent WO0226209.
ACCESSION AX468548
VERSION AX468548.1 GI:21901378
KEYWORDS
SOURCE Human immunodeficiency virus 1 (HIV-1)
ORGANISM Human immunodeficiency virus 1
VIRUSES; Retroviral viruses; Retroviridae; Lentivirus; Primate
Lentivirus group.
REFERENCE
1 O'Hagan,D., Otten,G., Donnelly,J.J., Polo,J.M., Barnett,S., Singh,M., Umer,J. and Dubensky,T.W. Microparticles for delivery of the heterologous nucleic acids Patent: WO 0226209-A 68 04-APR-2002;
JOURNAL CHIRON CORPORATION (US)
FEATURES
source location/Qualifiers
1. 1509
/organism="Human immunodeficiency virus 1"
/mol_type="unassigned DNA"
/db_xref="taxon:11676"
ORIGIN
Query Match 99.7%; Score 1504.2; DB 6; Length 1509;
Best Local Similarity 99.2%; Pred. No. 4.1e-150;
Matches 1497; Conservative 12; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGGCGCGCGCGCAGCATCTCTGCGCGCGGAGAGCTGSAACAAGTGGAGATCCGC 60
DB 1 ATGGGCGCGCGCGCAGCATCTCTGCGCGCGGAGAGCTGSAACAAGTGGAGATCCGC 60
QY 61 CTGCGCGCGCGCGCAGAGAGCACTACATGCTGAAGCACTGTGTGTGGCCAGCCGCGAG 120
DB 61 CTGCGCGCGCGCGCAGAGAGCACTACATGCTGAAGCACTGTGTGTGGCCAGCCGCGAG 120
QY 121 CTGAGAGGCTTCTGCTGTAACCTCCGCGCTGTGAGAGACCGCGGAGGGCTGCAAGCATC 180
DB 121 CTGAGAGGCTTCTGCTGTAACCTCCGCGCTGTGAGAGACCGCGGAGGGCTGCAAGCATC 180
QY 181 ATGAAGCAGCTGACAGCCCGCTGCAAGCCGCGACCGAGAGCTGCGCAGCTGTACAAC 240
DB 181 ATGAAGCAGCTGACAGCCCGCTGCAAGCCGCGACCGAGAGCTGCGCAGCTGTACAAC 240

```



```

DB 181 ATGAAGCAGCTGACCCCGCCTCTGACAGCCGACCCGAGAGCTCCGACGCTGTACAAC 240
QY 241 ACCGTGGCCACCTCTGTACTGCGTGCACGCGCGCATCGAGTCCGCGACACCAAGAGAGCC 300
DB 241 ACCGTGGCCACCTCTGTACTGCGTGCACGCGCGCATCGAGTCCGCGACACCAAGAGAGCC 300
QY 301 CTGGAACAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
DB 301 CTGGAACAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
QY 361 GCGGACGCGCAAGGTGAGCCAGAACTAACCCCATCGTGCAGAACTTGCAGAGGCGCAAGTGTG 420
DB 361 GCGGACGCGCAAGGTGAGCCAGAACTAACCCCATCGTGCAGAACTTGCAGAGGCGCAAGTGTG 420
QY 421 CACCAAGGCGCATGAGCCCGCGACACCTTGAAACGCTGGGTGAAGGTGATCGAGAGAGAGCC 480
DB 421 CACCAAGGCGCATGAGCCCGCGACACCTTGAAACGCTGGGTGAAGGTGATCGAGAGAGAGCC 480
QY 481 TTCAGCCCGGAGGTGATCCCGCATGTTCAACCGCCTGAGAGGCGCGCAACCCCGCAGAGAC 540
DB 481 TTCAGCCCGGAGGTGATCCCGCATGTTCAACCGCCTGAGAGGCGCGCAACCCCGCAGAGAC 540
QY 541 CTGAACAACGATGTTGAACAACCGTGGCGCGCGCAACGAGCGCGCATGAGATGCTGAAGAGAC 600
DB 541 CTGAACAACGATGTTGAACAACCGTGGCGCGCGCAACGAGCGCGCATGAGATGCTGAAGAGAC 600
QY 601 ACCATCAACGAGAGAGCGCGCGGAGTGGGAGACCGCTTGCAACCCCGTGCAGAGCGCGCGCTG 660
DB 601 ACCATCAACGAGAGAGCGCGCGGAGTGGGAGACCGCTTGCAACCCCGTGCAGAGCGCGCGCTG 660
QY 661 GCGCCCGCGCAGATGCGCGGAGACCCCGCGCGCGCGCGAGCATGCGCGCGCGCAACGAGACCTG 720
DB 661 GCGCCCGCGCAGATGCGCGGAGACCCCGCGCGCGCGCGAGCATGCGCGCGCGCAACGAGACCTG 720
QY 721 CAGAGAGAGATGCGCTGATGACCAAGCAACCCCGCGTGCACCGCGTGCAGATCTACAG 780
DB 721 CAGAGAGAGATGCGCTGATGACCAAGCAACCCCGCGTGCACCGCGTGCAGATCTACAG 780
QY 781 CCGTGGATCATCTGGGGCTGGAACAAGATGCTGCGGATGTACAGCCCGCTGAGCATCTG 840
DB 781 CCGTGGATCATCTGGGGCTGGAACAAGATGCTGCGGATGTACAGCCCGCTGAGCATCTG 840
QY 841 GACATCCGCGCAGGCGCGCGCAGAGAGCGCTTCCGCGACTACGTGAGCGCTTCTTCAAGACC 900
DB 841 GACATCCGCGCAGGCGCGCGCAGAGAGCGCTTCCGCGACTACGTGAGCGCTTCTTCAAGACC 900
QY 901 CTGCGCGCGCAGAGAGCGCGCACCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
DB 901 CTGCGCGCGCAGAGAGCGCGCACCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
QY 961 CAGAAACGCAACCCCGACTGCAAGACCATCTGCGCGCTCTCGCGCGCGCGCGCGCGCGCGCTG 1020
DB 961 CAGAAACGCAACCCCGACTGCAAGACCATCTGCGCGCTCTCGCGCGCGCGCGCGCGCGCGCTG 1020
QY 1021 GAGGAGATGATGACCGCTGCGCAGGCGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 1080
DB 1021 GAGGAGATGATGACCGCTGCGCAGGCGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 1080
QY 1081 GCGGAGCGGATGAGCGCGCGCAACAGCGGTGAACATCATGATGCAAGAGAGCAACTTCAAG 1140
DB 1081 GCGGAGCGGATGAGCGCGCGCAACAGCGGTGAACATCATGATGCAAGAGAGCAACTTCAAG 1140
QY 1141 GCGCGCGCGCGCAACGTCAGTGTCTTCAACTGCGCGCAAGAGAGGCGCAATCGCGCAAGAAC 1200
DB 1141 GCGCGCGCGCGCAACGTCAGTGTCTTCAACTGCGCGCAAGAGAGGCGCAATCGCGCAAGAAC 1200
QY 1201 TGCCTGCGCGCGCGCGCAAGAGAGGCTGCTGAAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
DB 1201 TGCCTGCGCGCGCGCGCGCAAGAGAGGCTGCTGAAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
QY 1261 GACTGCAACCGAGCGCGCAAGGCGCAACTTCTGCGCAAGATCTGCGCGCAAGAGGCGCGC 1320

```

```

DB 1261 GACTGCAACCGAGCGCGCAAGGCGCAACTTCTGCGCAAGATCTGCGCGCAAGAGGCGCGC 1320
QY 1321 CCGGCGCAACTTCTGTCAGAAACCGCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
DB 1321 CCGGCGCAACTTCTGTCAGAAACCGCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
QY 1381 CCGGCGGAGAGCTTCCGCTTTCAGAGAGAGACACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1440
DB 1381 CCGGCGGAGAGCTTCCGCTTTCAGAGAGAGACACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1440
QY 1441 CCGGAGCGCTTACCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1500
DB 1441 CCGGAGCGCTTACCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1500
QY 1501 AGCCAGTAA 1509
DB 1501 AGCCAGTAA 1509

RESULT 4
AX455905
LOCUS AX455905
DEFINITION Sequence 21 from Patent WO0204493.
ACCESSION AX455905
VERSION AX455905.1 GI:21714897
KEYWORDS
SOURCE
ORGANISM
synthetic construct
synthetic construct
artificial sequences.
REFERENCE
1 zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,B.
AUTHORS
TITL
JOURNAL
Patent: WO 0204493-A 21 17-JAN-2002;
CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)
FEATURES
source
1.1509
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="synthetic gag coding sequence of HIV strain
AF10967"

ORIGIN
Query Match 98.7%; Score 1489.8; DB 6; Length 1509;
Best Local Similarity 99.2%; Pred. No.1.3e-148;
Matches 1497; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 ATGGCGCGCGCGCGCAGCATCTGCGCGCGCGCGGAGAGCTGACCAAGTGGAGAGAGATCCGC 60
DB 1 ATGGCGCGCGCGCGCAGCATCTGCGCGCGCGCGGAGAGCTGACCAAGTGGAGAGAGATCCGC 60
QY 61 CTGCGCGCGCGCGCGCAGAGCACTACATGCTGAAGCACTGTGTGGGCCAGCGCGAG 120
DB 61 CTGCGCGCGCGCGCGCAGAGCACTACATGCTGAAGCACTGTGTGGGCCAGCGCGAG 120
QY 121 CTGAGGCGCTTGCCTGAAACCCCGCGCTGCGAGACCGCGAGGCGTGCAGAGCATC 180
DB 121 CTGAGGCGCTTGCCTGAAACCCCGCGCTGCGAGACCGCGAGGCGTGCAGAGCATC 180
QY 181 ATGAAGCAGTGCAGCGCGCGCGCTGCAACCGCGCAGAGAGCTGCGAGCGCTGTACAC 240
DB 181 ATGAAGCAGTGCAGCGCGCGCGCGCTGCAACCGCGCAGAGAGCTGCGAGCGCTGTACAC 240
QY 241 ACCGTGGCCACCTCTGTACTGCGTGCACGCGCGCATCGAGTCCGCGACACCAAGAGAGCC 300
DB 241 ACCGTGGCCACCTCTGTACTGCGTGCACGCGCGCATCGAGTCCGCGACACCAAGAGAGCC 300
QY 301 CTGGAACAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
DB 301 CTGGAACAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
QY 361 GCGGACGCGCAAGGTGAGCCAGAACTAACCCCATCGTGCAGAACTTGCAGAGGCGCAAGTGTG 420

```

```
Db 361 GCGGACGGCAAGGTGAGCCAGAACTAACCCTATGTCAGAACTTCAGAGGCGCCAGATGTG 420
Qy 421 CACCAAGGCAATCAAGCCCGCAGACCTTGAAAGCTGTGGTGAAGTGTATCGAGAGAGGCC 480
Db 421 CACCAAGGCAATCAAGCCCGCAGACCTTGAAAGCTGTGGTGAAGTGTATCGAGAGAGGCC 480
Qy 481 TTCAGCCCCGAGGTGATCCCCATGTTCAACGCGCTTGAGCGAGAGGCGCCACCCCCAGGAC 540
Db 481 TTCAGCCCCGAGGTGATCCCCATGTTCAACGCGCTTGAGCGAGAGGCGCCACCCCCAGGAC 540
Qy 541 CTGAACACGATGTTGAACACCGGTGGGCGGCGCAGAGGCGCGCATGCAAGTGTGAAGGAC 600
Db 541 CTGAACACGATGTTGAACACCGGTGGGCGGCGCAGAGGCGCGCATGCAAGTGTGAAGGAC 600
Qy 601 ACCATCAACGAGAGGCGCGCGAGTGGAGCCGCTTGACCCCGCTGCAAGGCGCGCCCGTG 660
Db 601 ACCATCAACGAGAGGCGCGCGAGTGGAGCCGCTTGACCCCGCTGCAAGGCGCGCCCGTG 660
Qy 661 GCGCGCGCGCAGATGCGCGAGACCCCGCGCGCAGCGACATGCGCGCGCGCCAGCAGCCTG 720
Db 661 GCGCGCGCGCAGATGCGCGAGACCCCGCGCGCAGCGACATGCGCGCGCGCCAGCAGCCTG 720
Qy 721 CAGGAGCAGATGCGCTGAGTGAACAGCAACCCCGCTGCGCGCGCGCGCATCTACAG 780
Db 721 CAGGAGCAGATGCGCTGAGTGAACAGCAACCCCGCTGCGCGCGCGCGCATCTACAG 780
Qy 781 CGGTGATCATCTTGGGCTGGAACAAGATCGTGGGATGTACAGCCCGTGAAGATCTG 840
Db 781 CGGTGATCATCTTGGGCTGGAACAAGATCGTGGGATGTACAGCCCGTGAAGATCTG 840
Qy 841 GACATCCGCGAGGCGCGCGCGCGCGCGCGCTTCCGCGACTAGTGAACCGCTTTCAGAGC 900
Db 841 GACATCCGCGAGGCGCGCGCGCGCGCGCGCTTCCGCGACTAGTGAACCGCTTTCAGAGC 900
Qy 901 CTGCGCGCGCAGAGGCGCGCGCGCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCG 960
Db 901 CTGCGCGCGCAGAGGCGCGCGCGCGCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCG 960
Qy 961 CAGAACGCGCAACCCCGACTGCAAGACCATCTGCGCGCGCGCGCGCGCGCGCGCGCG 1020
Db 961 CAGAACGCGCAACCCCGACTGCAAGACCATCTGCGCGCGCGCGCGCGCGCGCGCGCG 1020
Qy 1021 GAGGAGATGATGACCGCTGCGAGGCGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCG 1080
Db 1021 GAGGAGATGATGACCGCTGCGAGGCGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCG 1080
Qy 1081 GCGGAGGCGATGAGCCAGGCGCAAGCGTGAACATCATGATGCAAGAGCACTTCAG 1140
Db 1081 GCGGAGGCGATGAGCCAGGCGCAAGCGTGAACATCATGATGCAAGAGCACTTCAG 1140
Qy 1141 GCGCGCGCGCGCAACGTCAAGTCTTCAACTGCGCGCAAGGAGGCGCCATCGCCAAGAC 1200
Db 1141 GCGCGCGCGCGCAACGTCAAGTCTTCAACTGCGCGCAAGGAGGCGCCATCGCCAAGAC 1200
Qy 1201 TGCGCGCGCGCGCGCAAGAGGCGTGTGAGTGTGCGCAAGAGGCGCGCAAGATGAG 1260
Db 1201 TGCGCGCGCGCGCGCAAGAGGCGTGTGAGTGTGCGCAAGAGGCGCGCAAGATGAG 1260
Qy 1261 GACTGCAAGAGCGCGCAAGGCGCACTTCTGCGCAAGATCTGCGCGCAAGAGGCGCG 1320
Db 1261 GACTGCAAGAGCGCGCAAGGCGCACTTCTGCGCAAGATCTGCGCGCAAGAGGCGCG 1320
Qy 1321 CCGGCGCACTTCTGCGCAAGAGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
Db 1321 CCGGCGCACTTCTGCGCAAGAGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
Qy 1381 CCGGCGGAGGCTTCCGCTTGAAGAGACACCCCGCGCGCGCGCGCGCGCGCGCGCG 1440
Db 1381 CCGGCGGAGGCTTCCGCTTGAAGAGACACCCCGCGCGCGCGCGCGCGCGCGCGCG 1440
Qy 1441 CCGGAGCCCTACCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1500
Db 1441 CCGGAGCCCTACCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1500
```

```
Db 1441 CCGGAGCCCTACCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1500
Qy 1501 AGCCAGTAA 1509
Db 1501 AGCCAGTAA 1509

RESULT 5
AX455935
LOCUS AX455935
DEFINITION Sequence 51 from Patent WO0204493.
ACCESSION AX455935
VERSION AX455935.1 GI:21714919
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS zur Megede, J., Barnett, S.W., Engelbrecht, S. and van Rensburg, E.
TITLE Polynucleotides encoding antigenic hiv type c polypeptides,
JOURNAL Patent: WO 0204493-A 51 17-JAN-2002;
FEATURES
source
1. 1494
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="HIV Type C Gag optimized"

Query Match 86.0%; Score 1297.6; DB 6; Length 1494;
Best Local Similarity 92.7%; Pred. No. 2.6e-128;
Matches 1402; Conservative 0; Mismatches 89; Indels 21; Gaps 3;

Qy 1 ATGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
Db 1 ATGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
Qy 61 CTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
Db 61 CTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
Qy 121 CTGGAAGGCGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
Db 121 CTGGAAGGCGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
Qy 181 ATGAAGCAGCTGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
Db 181 ATGAAGCAGCTGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
Qy 241 ACCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
Db 241 ACCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
Qy 301 CTGGAACAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
Db 301 CTGGAACAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
Qy 361 GCGGAGCGGCAAGGTGAGCGCAAGCTAACCCCATGTGTGCAAGAACCTGCAAGGCG 420
Db 361 GCGGAGCGGCAAGGTGAGCGCAAGCTAACCCCATGTGTGCAAGAACCTGCAAGGCG 420
Qy 421 CACCAAGGCGCATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
Db 421 CACCAAGGCGCATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
Qy 481 TTCAGCCCCGAGGTGATCCCCATGTTCAACGCGCTTGAGCGAGAGGCGCCACCCCG 540
Db 481 TTCAGCCCCGAGGTGATCCCCATGTTCAACGCGCTTGAGCGAGAGGCGCCACCCCG 540
Qy 541 CTGAACACGATGTTGAACACCGGTGGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
Db 541 CTGAACACGATGTTGAACACCGGTGGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
```

```

Db      541 CTGAACACCACTGTAACACCGTGGGCGGCAACGAGCCGCAATGAGATGCTGAAGAC 600
QY      601 ACCATCAACGAGAGGCGCGGAGTGGGACCGCCTGCAACCCGTCAGAGCCGCGCCGCTG
      601 ACCATCAACGAGAGGCGCGGAGTGGGACCGCAGCCACCCCGTCAGAGCCGCGCCGCTG
QY      661 GCGCGCGCGGAGTGGGCGGAGCGCGCGGAGCATGCGCGCGGCAACGAGACCCCTG
      661 GCGCGCGCGGAGTGGGCGGAGCGCGCGGAGCATGCGCGCGGCAACGAGACCCCTG
Db      721 CAGAGGAGATGCGCTGATGACGAGCAACCCCGCTGCGCGGAGCATCTACAAG 780
QY      721 CAGAGGAGATGCGCTGATGACGAGCAACCCCGCTGCGCGGAGCATCTACAAG 780
Db      781 CGGTGATCATCTGCGGCTGTAACAAGATCGTGGCATGTACAGCCCGTGAGCATCTG 840
QY      781 CGGTGATCATCTGCGGCTGTAACAAGATCGTGGCATGTACAGCCCGTGAGCATCTG 840
Db      841 GACATCCGCGAGGCGCGGAGGAGCGCTTCCGCGACTACGTGAGCCGCTTCTCAAGACC 900
QY      841 GACATCAAGGAGGCGCGGAGGAGCGCTTCCGCGACTACGTGAGCCGCTTCTCAAGACC 900
Db      901 CTGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960
QY      901 CTGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960
Db      961 CAGAACGCGCAACCCCGACTGCAAGACATCTGCGCGCTCTCGCGCGCGGAGCCCTG 1020
QY      961 CAGAACGCGCAACCCCGACTGCAAGACATCTGCGCGCTCTCGCGCGCGGAGCCCTG 1020
Db      1021 GAGGAGATGATGACCGCTGCGAGGCGGTGGGCGCGCGGAGGAGGAGGAGGAGGAGGAG 1080
QY      1021 GAGGAGATGATGACCGCTGCGAGGCGGTGGGCGCGCGGAGGAGGAGGAGGAGGAG 1080
Db      1081 GCGGAGGCGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
QY      1081 GCGGAGGCGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
Db      1138 GCGAGCAACCGCATCATCAAGTCTCACTGCGGCAAGTGGGCAATGCGCGCAAC 1197
QY      1138 GCGAGCAACCGCATCATCAAGTCTCACTGCGGCAAGTGGGCAATGCGCGCAAC 1197
Db      1201 TGCCTGCGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
QY      1201 TGCCTGCGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
Db      1198 TGCCTGCGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1257
QY      1261 GACTGCAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
Db      1258 GACTGCAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1317
QY      1321 CCGCGCAACTTCTGCAAGACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1377
Db      1318 CCGCGCAACTTCTGCAAGACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1377
QY      1378 CCGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1437
Db      1378 CCGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1437
QY      1438 GACCGCGAGGCGCTACCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1497
Db      1438 -----CGCGAGCGCTGACAGGCTGAAGAAGCTGTCGCGAGCGAGCC 1482
QY      1498 CTGAGCCAGTAA 1509
Db      1483 CTGAGCCAGTAA 1494

```

```

RESULT 6
LOCUS   AX455983
DEFINITION Sequence 99 from Patent WO0204493.
ACCESSION AX455983
VERSION  AX455983.1 GI:21714967

```

```

KEYWORDS
SOURCE   synthetic construct
ORGANISM synthetic construct
REFERENCE
AUTHORS  zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,B.
TITLE    Polynucleotides encoding antigenic hiv type c polypeptides,
JOURNAL  Patent: WO 0204493-A 99 17-JAN-2002;
          CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)
FEATURES
         Source
         1..1491
         /organism="synthetic construct"
         /mol_type="unassigned DNA"
         /db_xref="taxon:32630"
         /note="Gag_TV2_C_ZAopt"
ORIGIN
Query Match 85.8%; Score 1294.8; DB 6; Length 1491;
Best Local Similarity 92.7%; Pred. No. 5,2e-128;
Matches 1401; Conservative 0; Mismatches 87; Indels 24; Gaps 3;

QY      1 ATGGGCGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 60
Db      1 ATGGGCGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 60
QY      61 CTGCGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
Db      61 CTGCGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
QY      121 CTGAGGCGCTTCCGCTGAAACCGCGGCTGCTGAGAGACCGCGGAGGAGGAGGAGGAG 180
Db      121 CTGAGGCGCTTCCGCTGAAACCGCGGCTGCTGAGAGACCGCGGAGGAGGAGGAGGAG 180
QY      181 ATGAAGCAGCTGAGACCGCGGCTGAGAGACCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
Db      181 ATGAAGCAGCTGAGACCGCGGCTGAGAGACCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
QY      241 ACCGTGGCAACCTGTACTGCTGTCAGCGCGGCAATGAGGTCGCGGACACCAAGAGGCC 300
Db      241 ACCGTGGCAACCTGTACTGCTGTCAGCGCGGCAATGAGGTCGCGGACACCAAGAGGCC 300
QY      301 CTGGAACAAGTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
Db      301 CTGGAACAAGTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
QY      361 GCGGAGGCAAGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
Db      361 GCGGAGGCAAGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
QY      421 CACGAGGCGATCAGCCCGCGACCTGGAAGCGCTGGTGAAGGTGATCGAGGAGGAGGAGGAG 480
Db      421 CACGAGGCGATCAGCCCGCGACCTGGAAGCGCTGGTGAAGGTGATCGAGGAGGAGGAGGAG 480
QY      481 TTGAGCCCGGAGGTGATCCCATGTTCAACCGCTGAGCGAGGAGGAGGAGGAGGAGGAGGAG 540
Db      481 TTGAGCCCGGAGGTGATCCCATGTTCAACCGCTGAGCGAGGAGGAGGAGGAGGAGGAGGAG 540
QY      541 CTGAACAGATGTTGAACAACGCTGGGCGGCAACGAGCGGCGCATGAGATGCTGAAGAGC 600
Db      541 CTGAACAGATGTTGAACAACGCTGGGCGGCAACGAGCGGCGCATGAGATGCTGAAGAGC 600
QY      601 ACCATCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
Db      601 ACCATCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
QY      661 GCGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
Db      661 GCGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
QY      721 CAGGAGCAGATGCGCTGATGAGCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
Db      721 CAGGAGCAGATGCGCTGATGAGCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780

```





```

OY 535 CAGGACCTGAACAAGATGTTGAACACCTGCGGCGCCACGAGCCGCGCATGAGATGCTG 594
    |||||
DB 541 CAGGACCTGAACAAGATGTTGAACACCTGCGGCGCCACGAGCCGCGCATGAGATGCTG 600
OY 595 AAGGACACATCAACGAGGAGCGCGGAGTGGAGCCGCTTGCACCCCGTGACGGCCGCG 654
    |||||
DB 601 AAGGACACATCAACGAGGAGCGCGGAGTGGAGCCGCTTGCACCCCGTGACGGCCGCG 660
OY 655 CCCGTGCGCGCGCGGAGATGCGGAGCCCGCGCGGAGGAGATGCGCGCGCGCGCGCGAGC 714
    |||||
DB 661 CCCGTGCGCGCGCGGAGATGCGGAGCCCGCGCGGAGGAGATGCGCGCGCGCGCGAGC 720
OY 715 ACCCTGCAAGAGCAGATGCGCTGAGTACCAAGACCCCGCGTGCCTGCGGCGAGCATC 774
    |||||
DB 721 ACCCTGCAAGAGCAGATGCGCTGAGTACCAAGACCCCGCGTGCCTGCGGCGAGCATC 780
OY 775 TACAAGCGGTGATCATCTGCGGCTGAAACAAGATGTCGCGATGTACAGCCCGTGAGC 834
    |||||
DB 781 TACAAGCGGTGATCATCTGCGGCTGAAACAAGATGTCGCGATGTACAGCCCGTGAGC 840
OY 835 ATCCCTGCAATCCGCGAGGCGCGCAAGAGCCCTTCCGCGACTACGTGAGCCGCTTCTC 894
    |||||
DB 841 ATCCCTGCAATCAAGCAGGCGCGCAAGAGCCCTTCCGCGACTACGTGAGCCGCTTCTC 900
OY 895 AAGACCTGCGCGCGCGGAGCAGGCGCAACCAAGAGTGAAGACTGATGACCGACACCTG 954
    |||||
DB 901 AAGACCTGCGCGCGCGGAGCAGGCGCAACCAAGAGTGAAGACTGATGACCGACACCTG 960
OY 955 CTGCTGCAAGACCGCAACCGCGCACTGCAAGACCATCTGCGCGCTTCCGCGCGCGC 1014
    |||||
DB 961 CTGCTGCAAGACCGCAACCGCGCACTGCAAGACCATCTGCGCGCTTCCGCGCGCGC 1020
OY 1015 ACCCTGAGAGATGATGACCGCTGCGAGGCGTGGCGCGCGCGCGCGCGCGCGCGCGC 1074
    |||||
DB 1021 ACCCTGAGAGATGATGACCGCTGCGAGGCGTGGCGCGCGCGCGCGCGCGCGCGCGC 1080
OY 1075 GTGCTGCGCGCGCGGAGTGAAGCGCAACGCTGGAATCATGATGCAAGAGCAAC 1134
    |||||
DB 1081 GTGCTGCGCGCGCGGAGTGAAGCGCAACGCTGGAATCATGATGCAAGAGCAAC 1140
OY 1135 TTCAAGGCGCGCGCGCGCAACGCTCAAGTCTTCACTGCGCGCAAGAGGCGCGCATGCGC 1194
    |||||
DB 1141 TTCAAGGCGCGCGCGCGCAACGCTCAAGTCTTCACTGCGCGCAAGAGGCGCGCATGCGC 1200
OY 1195 AAGACTGCGCGCGCGCGCGCAAGAGGCTGCTGAAGTGGCGCAAGAGGCGCGCACG 1254
    |||||
DB 1201 CGCAACTGCGCGCGCGCGCGCGCAAGAGGCTGCTGAAGTGGCGCAAGAGGCGCGCACG 1260
OY 1255 ATGAAGACTGCAACGAGCGCGCAACTTCTGCGCAAGATCTGCGCGCACGCAAG 1314
    |||||
DB 1261 ATGAAGACTGCAACGAGCGCGCAACTTCTGCGCAAGATCTGCGCGCACGCAAG 1320
OY 1315 GCGCGCGCGCGCACTTCTGCAAGACGCGAGCGCGCGCGCGCGCGCGCGCGCGCACG 1374
    |||||
DB 1321 GCGCGCGCGCGCACTTCTGCAAGACG-----CGCGAGCGCGCACG 1362
OY 1375 GCGCGCGCGCGCGAGAGCTTCCGCTTGAAGAGCAACCGCGCGCGCGCGCGCGCGCGCGC 1434
    |||||
DB 1363 GCGCGCGCGCGCGAGAGCTTCCGCTTGAAGAGCAACCGCGCGCGCGCGCGCGCGCGCGC 1422
OY 1435 AAGGACCGGAGCGCTTACCGCGAGCGCGCTGACCGCGCTGCGAGCTGTCGCGAGCGC 1494
    |||||
DB 1423 AAGG-----ACCGCGAGCGCGCTGACCGAGCTTGAAGAGCTGTCGCGAGCGAGC 1470
OY 1495 CCCCTGAGCGAGTAA 1509
    |||||
DB 1471 CCCCTGAGCGAGTAA 1485

```

```

RBSULT 8
LOCUS AX455887 1479 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 3 from Patent WO0204493.

```

```

ACCESSION AX455887
VERSION AX455887.1 GI:21714880
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS zur Mege, J., Barnett, S.W., Engelbrecht, S. and van Rensburg, E.
TITLE Polynucleotides encoding antigenic hiv type c polypeptides,
JOURNAL Patent: WO 0204493-A 3 17-JAN-2002;
CHIRON CORPORATION (US); University of Stellenbosch (ZA)
FEATURES
source
1. 1479
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="synthetic Gag of HIV strain AF110965"
ORIGIN

```

```

Query Match 84.6%; Score 1276.8; DB 6; Length 1479;
Best Local Similarity 92.5%; Pred. No. 4,1e-126;
Matches 1399; Conservative 0; Mismatches 77; Indels 36; Gaps 4;

```

```

OY 1 ATGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
    |||||
DB 1 ATGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
OY 61 CTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
    |||||
DB 61 CTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
OY 121 CTGAGGCGCTTCCGCTGAACTCCGCGCGCTGCTGGAAGACCGCGCGCGCGCGCGCGCG 180
    |||||
DB 121 CTGAGGCGCTTCCGCTGAACTCCGCGCGCTGCTGGAAGACCGCGCGCGCGCGCGCGCG 180
OY 181 ATGAAGCAGCTGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
    |||||
DB 181 ATCGCGCAGCTGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
OY 241 ACCGTGCGCAACCTGTACTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
    |||||
DB 241 ACCGTGCGCAACCTGTACTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
OY 301 CTGGAACAAGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
    |||||
DB 301 CTGGAACAAGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
OY 361 GCGGAC--GCGAAGGTGAGCGCAAGTACCCCATGCTGTCAGAACTTGCAGGCGCGCATG 417
    |||||
DB 361 GCGGAC--GCGAAGGTGAGCGCAAGTACCCCATGCTGTCAGAACTTGCAGGCGCGCATG 420
OY 418 GTGCAACGAGCGCATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 477
    |||||
DB 421 GTGCAACGAGCGCATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
OY 478 GCGTTCAGCGCGCGAGGTGATCCCATGTTCAACCGCGCTGAGCGAGGCGCGCGCGCGCG 537
    |||||
DB 481 GCGTTCAGCGCGCGAGGTGATCCCATGTTCAACCGCGCTGAGCGAGGCGCGCGCGCGCG 540
OY 538 GACCTGAACAAGATGTTGAACCGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 597
    |||||
DB 541 GACCTGAACAAGATGTTGAACCGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
OY 598 GACACCATCAACGAGAGGCGCGCGAGTGGAGCCGCTGCAACCGCGCTGCAAGCGCGCGCG 657
    |||||
DB 601 GACACCATCAACGAGAGGCGCGCGAGTGGAGCCGCTGCAACCGCGCTGCAAGCGCGCGCG 660
OY 658 GTGCGCGCGCGCGCGAGTGGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 717
    |||||
DB 661 ATGCGCGCGCGCGCGAGTGGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
OY 718 CTGAGGAGCAGATGCGCTGATGACAGCAACCGCGCGTGCCTGCGCGCGCGCGCGCATCTAC 777

```



```

Db      721 CTGACAGACAGATCGCTGTGATGACAGCAACCCCCCATCCCGTGGGCAATCTAC 780
Qy      778 AAGCGGTGATCATCTTGCGCTGAACAAGATGTCGGATGTAAGCCCGGTGAGCATC 837
Db      781 AAGCGGTGATCATCTTGCGCTGAACAAGATGTCGGATGTAAGCCCGGTGAGCATC 840
Qy      838 CTGACATCCCGCAGGCGCCCAAGAGCCCTTCGCGACTACGTGACCGCTTCTCAAG 897
Db      841 CTGACATCAAGCAGGCGCCCAAGAGCCCTTCGCGACTACGTGACCGCTTCTCAAG 900
Qy      898 ACCCTGCGCGCGAGCAGAGCAACCCAGAGCTGAAGAACTGATGACCGAGACCTGCTG 957
Db      901 ACCCTGCGCGCGAGCAGAGCAACCCAGAGCTGAAGAACTGATGACCGAGACCTGCTG 960
Qy      958 GTGAGAAAGCAACCCCGCACTGCAAGACCACTCTGCGCTCTGCGCCCGCGCCAGC 1017
Db      961 GTGAGAAAGCAACCCCGCACTGCAAGACCACTCTGCGCTCTGCGCCCGCGCCAGC 1020
Qy      1018 CTGAGAGATGATGACCGCTGCGAGGCGTGGCGCGCCCGCCAGAGCGCGCTG 1077
Db      1021 CTGAGAGATGATGACCGCTGCGAGGCGTGGCGCGCCCGCCAGAGCGCGCTG 1080
Qy      1078 CTGCGCGAGCGATGAGCGCAAGCGCAACAGCTGAACATCATGATGAGAGCAACTTC 1137
Db      1081 CTGCGCGAGCGATGAGCGCAAGCGCAACAGCTGAACATCATGATGAGAGCAACTTC 1137
Qy      1138 AAGGCGCGCGCGCAACCTCAAGTCTTCACTGCGCAAGAGAGGCGCAATCGCCAG 1197
Db      1138 AAGGCGCGCGCGCAATCTCAAGTCTTCACTGCGCAAGAGAGGCGCAATCGCCAG 1197
Qy      1198 AACTGCGCGCGCGCGCGCAAGAGGCTGCTGAAGTGGCGCAAGAGGCGCCACAGATG 1257
Db      1198 AACTGCGCGCGCGCGCGCAAGAGGCTGCTGAAGTGGCGCAAGAGGCGCCACAGATG 1257
Qy      1258 AAGGACTGACCGAGCGCGCAAGCTTCTGCGCAAGATCTGCGCCAGCAAGAGGCG 1317
Db      1258 AAGGACTGACCGAGCGCGCAAGCTTCTGCGCAAGATCTGCGCCAGCAAGAGGCG 1317
Qy      1318 CGCGCGCGCACTTCTGCAAGACCGAGAGCGCGCGCGCCCAACGCTGCCACCGCC 1377
Db      1318 CGCGCGCGCACTTCTGCAAGACCGAGAGCGCGCGCGCCCAACGCTGCCACCGCC 1377
Qy      1378 CCCCCCGCAGAGCTTCCGCTTGAAGAGCAACCCCGCGCCCAAGCAGAGCCCAAG 1437
Db      1360 CCCCCCGCAGAGCTTCCGCTTGAAGAGCAACCCCGCGCCCAAGCAGAGCCCAAG 1419
Qy      1438 GACCGCGAGCCCTACCGCGAGCCCTGACCGCCCTGCGAGCCTGTTGGCAGCGGCC 1497
Db      1420 G-----ACCGCGAGACCTTGAACAGCCTGAAGAGCCTGTTGGCAGAGCC 1467
Qy      1498 CTGAGCCAGTAA 1509
Db      1468 CTGAGCCAGTAA 1479

```

```

RESULT 9
AX468543          1479 bp      DNA      linear      PAT 16-JUL-2002
LOCUS             AX468543
DEFINITION       Sequence 63 from Patent WO226209.
ACCESSION        AX468543
VERSION          AX468543.1 GI:21901373
KEYWORDS
SOURCE           Human immunodeficiency virus 1 (HIV-1)
ORGANISM         Human immunodeficiency virus 1
                 Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
                 lentivirus group.
REFERENCE
AUTHORS          1 O'Hagan,D., Otten,G., Donnelly,J.J., Polo,J.M., Barnett,S.,
                  Singh,M., Uimer,J. and Dubensky,T.W.
TITLE            Microparticles for delivery of the heterologous nucleic acids
JOURNAL          Patent: WO 0226209-A 63 04-APR-2002;
                  CHIRON CORPORATION (US)

```

```

FEATURES
    source
        Location/Qualifiers
            1..1479
            /organism="Human immunodeficiency virus 1"
            /mol_type="unassigned DNA"
            /db_xref="taxon:11676"
    ORIGIN
        Query Match      84.6%; Score 1276.8; DB 6; Length 1479;
        Best Local Similarity 92.5%; Pred. No. 4,1e-126;
        Matches 1399; Conservative 0; Mismatches 77; Indels 36; Gaps 4;
Qy      1 ATGGCGCGCGCGCAGCATCTGCGCGCGCAAGCTGAGCAAGTGGAGAGATCCGC 60.
Db      1 ATGGCGCGCGCGCAGCATCTGCGCGCGCAAGCTGAGCATGAGCGCTGGAGCGCATCCGC 60
Qy      61 CTGCGCGCGCGCGCAAGAGCACTACATGCTGAAGCACTGTGTGGCCAGCGCGAG 120
Db      61 CTGCGCGCGCGCGCAAGAGTGTACATGATGAAGCACTGTGTGGCCAGCGCGAG 120
Qy      121 CTGAGAGCTTCCCTGAAACCCCGCGCTGTGAGAGACCGCGAGGGCTGCAAGCATC 180
Db      121 CTGAGAGTTCCTGCTGAACCCCGCGCTGTGAGAGACCGAGGGCTGCAAGCATC 180
Qy      181 ATGAAGCAGCTGACCGCGCGCTGACAGCGCGCAACGAGAGCTGCGAGCTGTCAAC 240
Db      181 ATCGCCAGCTGACCGCGCGCTGACAGCGCGCAACGAGAGCTGAGAGCTGTTCAC 240
Qy      241 ACCGTGCGCACTGTACTGCTGTGACGCGCGCATGAGGTCCGCAACCAAGAGGCC 300
Db      241 ACCGTGCGCACTGTACTGCTGTGACGAGAGAGATGAGGTCCGCAACCAAGAGGCC 300
Qy      301 CTGACAAGATCGAGAGAGCAAGCAAGTCCAGCAAGAGACCGAGCGCAAGAG 360
Db      301 CTGACAAGATCGAGAGAGCAAGCAAGTCCAGCAAGAGATCCAGAGCGCGAGGCC 360
Qy      361 GCCGAC---GCAAGGTGAGCGCAAGACTACCCCATGCTGTGAGAACTGAGGCGCAGATG 417
Db      361 GCCGACAAGGCGCAAGGTGAGCGCAAGACTACCCCATGCTGTGAGAACTGAGGCGCAGATG 420
Qy      418 GTGACACAGGCGCATGACCGCGCAACCTGAACGCTGGGTGAAGGTGATGAGAGAAG 477
Db      421 GTGACACAGGCGCATGACCGCGCAACCTGAACGCTGGGTGAAGGTGATGAGAGAAG 480
Qy      478 GCCTTCAAGCGCGAGGTGATCCCATGTTCAACCGCCCTGAGCGAGGGCGCACCCCGAG 537
Db      481 GCCTTCAAGCGCGAGGTGATCCCATGTTCAACCGCCCTGAGCGAGGGCGCACCCCGAG 540
Qy      538 GACCTGAACAGATGTTGAACACCGTGGCGCGCACAGGCGCGCATGAGATCTGAAG 597
Db      541 GACCTGAACAGATGTTGAACACCGTGGCGCGCACAGGCGCGCATGAGATCTGAAG 600
Qy      598 GACACCATCAACGAGAGGCGCGAGTGGACCGGCTGACCCCGTGAGGCGGCGCC 657
Db      601 GACACCATCAACGAGAGGCGCGAGTGGACCGGCTGACCCCGTGACCGGCGGCGCC 660
Qy      658 GTGCGCGCGCGCAGATGCGCGACCGCGCGCGCGAGCATGCGCGCGCGCACAGCAC 717
Db      661 ATCGCGCGCGCGCAGATGCGCGAGCGCGCGCGCGAGCATGCGCGCGCGCACAGCAC 720
Qy      718 CTGCAAGAGCAGATGCGCTGATGAACCAAGCAACCCCGCTGCGCGTGGCGCATCTAC 777
Db      721 CTGCAAGAGCAGATGCGCTGATGAACCAAGCAACCCCGCATCCCGTGGCGCATCTAC 780
Qy      778 AAGCGGTGATCATCTGCGCTGAACAAGATGTCGGATGTAAGCCCGGTGAGCATC 837
Db      781 AAGCGGTGATCATCTGCGCTGAACAAGATGTCGGATGTAAGCCCGGTGAGCATC 840
Qy      838 CTGACATCCCGCAGGCGCCCAAGAGCGCTTCCGCACTACCTGGAACCGCTTCTCAAG 897
Db      841 CTGACATCAAGCAGGCGCCCAAGAGCGCTTCCGCACTACCTGGAACCGCTTCTCAAG 900
Qy      898 ACCCTGCGCGCGAGCGCAAGCGCAAGAGCTGAAGAACTGATGACCGAGACCTGCTG 957

```

Db 901 ACCCTGCGCGCCGAGCAGAGCAACCAAGAGTGAAGAACTGATGACCGACACCTGCTG 960  
QY 958 GTGCAAGACGCGCAACCCCGACTGCAAGACCATCTGCGCGCTCTGCGCCCGCGCGCAC 1017  
Db 961 GTGCAAGACGCGCAACCCCGACTGCAAGACCATCTGCGCGCTCTGCGCCCGCGCGCAC 1020  
QY 1018 CTGAGAGAGATGATGACCGCTGCGAGAGGCGTGGCGGCGCCCGCGCAAGGCGCGCTG 1077  
Db 1021 CTGAGAGAGATGATGACCGCTGCGAGAGGCGTGGCGGCGCCCGCGCAAGGCGCGCTG 1080  
QY 1078 CTGAGCGAGCGATGAGCGCAGGCGCAACAGCGTGAACATCATGATGAGAGAGCACTTC 1137  
Db 1081 CTGAGCGAGCGATGAGCGCAGGCGCAACAGCGTGAACATCATGATGAGAGAGCACTTC 1137  
QY 1138 AAGGCGCGCGCGCGCAACGTCAAGTCTTCAACTGCGCGCAAGAGGCGCACATGCGCAAG 1197  
Db 1138 AAGGCGCGCGCGCGCGCAACGTCAAGTCTTCAACTGCGCGCAAGAGGCGCACATGCGCGC 1197  
QY 1198 AACTGCGCGCGCGCGCGCAAGAGGCGTGTGAAGTGGCGCAAGAGGCGCGCACAGATG 1257  
Db 1198 AACTGCGCGCGCGCGCGCGCAAGAGGCGTGTGAAGTGGCGCAAGAGGCGCGCACAGATG 1257  
QY 1258 AAGGACTGCAACCGAGCGCGCAACTTCTGCGCAAGATCTGCGCGCAAGCGCACAGGCG 1317  
Db 1258 AAGGACTGCAACCGAGCGCGCGCAACTTCTGCGCGCAAGATCTGCGCGCAAGCGCGCGC 1317  
QY 1318 CGCGCGCGCGCAACTTCTGCGCAAGAGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCG 1377  
Db 1318 CGCGCGCGCGCAACTTCTGCGCAAGAGCGCG-----CGCGAGCGCGCGCGCGCGCG 1359  
QY 1378 CGCGCGCGCGAGCGTCTGCGTTCGAGAGAGACCAACCGCGCGCGCGCGCGCGCGCGCG 1437  
Db 1360 CGCGCGCGCGAGCGTCTGCGTTCGAGAGAGACCAACCGCGCGCGCGCGCGCGCGCGCG 1419  
QY 1438 GACCGCGAGCGCTTACCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1497  
Db 1420 G-----ACCGGAGAGCGCTGAGCGAGCGCTGTAAGAGCGCTGTTGCGCAAGCGCG 1467  
QY 1498 CTGAGCCAGTAA 1509  
Db 1468 CTGAGCCAGTAA 1479

RESULT 10  
AX468547 1479 bp DNA linear PAT 16-JUL-2002  
LOCUS AX468547  
DEFINITION Sequence 67 from Patent WO0226209.  
ACCESSION AX468547  
VERSION AX468547.1 GI:21901377

KEYWORDS Human immunodeficiency virus 1 (HIV-1)  
SOURCE Human immunodeficiency virus 1  
ORGANISM Human immunodeficiency virus 1  
VIRUSES; Retroviral viruses; Retroviridae; Lentivirus; Primate  
Lentivirus group.

REFERENCE  
AUTHORS O'Hagan, D., Otten, G., Donnelly, J. J., Polo, J. M., Barnett, S.,  
Singh, M., Ulmer, J. and Dubensky, T. W.  
TITLE Microparticles for delivery of the heterologous nucleic acids  
JOURNAL Patent: WO 0226209-A 67 04-APR-2002;  
CHIRON CORPORATION (US)  
FEATURES Location/Qualifiers

source 1.1479  
/organism="Human immunodeficiency virus 1"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:11676"

## ORIGIN

Query Match 84.2%; Score 1271.2; DB 6; Length 1479;  
Best Local Similarity 91.8%; Pred. No. 1.6e-125;  
Matches 1388; Conservative 10; Mismatches 78; Indels 36; Gaps 4;  
QY 1 ATGGGCG 60  
|||||  
Db 1 ATGGGCG 60  
|||||

Db 1 ATGGGCG 60  
QY 61 CTGCG 120  
Db 61 CTGCG 120  
QY 121 CTGAGAGGCTTGGCGCTGAAACCGCGCGCTGCTGAGAACCGCGAGGCTGCAAGCAATC 180  
Db 121 CTGAGAGGCTTGGCGCTGAAACCGCGCGCTGCTGAGAACCGCGAGGCTGCAAGCAATC 180  
QY 181 ATGAAGCAGCTGACCG 240  
Db 181 ATCCG 240  
QY 241 ACCGTGGCG 300  
Db 241 ACCGTGGCG 300  
QY 301 CTGGAACAAGATCGAG 360  
Db 301 CTGGAACAAGATCGAG 360  
QY 361 GCGGAC--GGCAAGGTGAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 417  
Db 361 GCGGACAG 420  
QY 418 GTGCAACAAGCG 477  
Db 421 GTGCAACAAGCG 480  
QY 478 GCCTTCAGCG 537  
Db 481 GCCTTCAGCG 540  
QY 538 GACCTGAACAGGATGTTGAACACCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 597  
Db 541 GACCTGAACAGGATGTTGAACACCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600  
QY 598 GACACCATCAACGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 657  
Db 601 GACACCATCAACGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660  
QY 658 GTGCG 717  
Db 661 ATGCG 720  
QY 718 CTGAGAGAGAGAGATGCGCTGAGTGAACAGCAACCGCGCGCGCGCGCGCGCGCG 777  
Db 721 CTGAGAGAGAGAGATGCGCTGAGTGAACAGCAACCGCGCGCGCGCGCGCGCGCG 780  
QY 778 AAGCGGTGATCATCTGCGCGCTGGAACAAGATCTGCGCGATGTACAGCGCGTGAGATC 837  
Db 781 AAGCGGTGATCATCTGCGCGCTGGAACAAGATCTGCGCGATGTACAGCGCGTGAGATC 840  
QY 838 CTGACATTCGCGAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 897  
Db 841 CTGACATTCGCGAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900  
QY 898 ACCGTGCGCGCGCGAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 957  
Db 901 ACCGTGCGCGCGCGAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 960  
QY 958 GTGCAAGACGCGCAACCGCGACTGCAAGACCATCTGCGCGCTCTGCGCGCGCGCGCAC 1017  
Db 961 GTGCAAGACGCGCAACCGCGACTGCAAGACCATCTGCGCGCTCTGCGCGCGCGCGCAC 1020  
QY 1018 CTGAGAGAGATGATGACCGCTGCGAGAGGCGTGGCGGCGCCCGCGCAAGGCGCGCTG 1077  
Db 1021 CTGAGAGAGATGATGACCGCTGCGAGAGGCGTGGCGGCGCCCGCGCAAGGCGCGCTG 1080  
QY 1078 CTGCGCGAGCGATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1137  
Db 1081 CTGCGCGAGCGATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1137



QY 1318 CGCCCCGGCACTTCTGCAAGAACCGAGCCCGCCGCCACCGTCCCAACGCC 1377  
DB 1318 CGCCCCGGCACTTCTGCAAGACG-----CCCCGAGCCCAACGCC 1359  
QY 1378 CCCCCCGCGAGAGCTTCCGCTTGAGAGAACACCCCGCCCAAGCAGAGCCCAAG 1437  
DB 1360 CCCCCCGCGAGAGCTTCCGCTTGAGAGAACACCCCGCCCAAGCAGAGCCCAAG 1419  
QY 1438 GACCGCGAGCCCTACCGCGAGCCCTGACCGCCCTGCGCAGCCTGTTCCGCGAGCGCC 1497  
DB 1420 G-----ACCGCGAGACCTTGACCGCCTGAAGAGCCTGTTCCGCAACGACCCC 1467  
QY 1498 CTGAGCCAGTAA 1509  
DB 1468 CTGAGCCAGTAA 1479

RESULT 12  
LOCUS AX149648 4288 bp DNA linear PAT 08-JUN-2001  
DEFINITION Sequence 2 from Patent WO0136614.  
ACCESSION AX149648  
VERSION AX149648.1 GI:14348047

KEYWORDS Human immunodeficiency virus  
SOURCE Human immunodeficiency virus  
ORGANISM Viruses; Retroviridae; Lentivirus; Primate  
lentivirus group.

REFERENCE  
1 Shao, Y., Wagner, R., Wolf, H. and Graf, M.  
AUTHORS The genome of the hiv-1 inter-subtype (c/b') and use thereof  
TITLE Patent: WO 0136614-A 2 25-MAY-2001;  
JOURNAL Genent GMBH Gesellschaft fuer angewandte Boitechnologie (DE) ;  
Sha, Yiming (CN)  
FEATURES  
LOCATION/Qualifiers  
1..4288  
Source /organism="Human immunodeficiency virus"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:12721"

ORIGIN  
Query Match 79.7%; Score 1202.2; DB 6; Length 4288;  
Best Local Similarity 88.5%; Pred. No. 2.3e-118;  
Matches 1335; Conservative 0; Mismatches 153; Indels 21; Gaps 2;

QY 1 ATGGGCGCGCGCGCGAGATCTGCGCGCGAGAGCTGAGCAAGTGGAGAGATCCGC 60  
DB 13 ATGGGCGCGCGCGCGAGATCTGAGGCGCGAGAGCTGAGCAAGTGGAGAGATCAGG 72  
QY 61 CTGCGCGCGCGCGCGAGAGCACTACATGCTGAGCAACCTGTGTGGCGAGCGCGAG 120  
DB 73 CTGAGGCGCGCGCGCGAGAGCACTACATGCTGAGCAACCTGTGTGGCGAGCGAG 132  
QY 121 CTGAGAGGCTTCCGCTTGAACCCCGCGCTGCTGAGAGCCCGAGGGCTGCAAGCATC 180  
DB 133 CTGAGAGGCTTCCGCTTGAACCCCGCGCTGCTGAGAGCCCGAGGGCTGCAAGCATC 192  
QY 181 ATGAAGCAGCTGCAAGCCCGCGCTGCAAGCCCGCAACGAGAGCTGCGCAGCTGTAAAC 240  
DB 193 ATGAAGCAGCTGCAAGCCCGCGCTGCAAGCCCGCAACGAGAGCTGCGCAGCTGTAAAC 252  
QY 241 ACCGTGCGCACTCTGTACTGCTGTGCAAGCCCGCATGAGGTCCGCGACACCAAGAGGCC 300  
DB 253 ACCGTGCGCACTCTGTACTGCTGTGCAAGCCCGCATGAGGTCCGCGAGAGGCC 312  
QY 301 CTGCAACAAGTCAAGAGAGAGCAACAAGTCCAGCAAGAACCCAGAGGCCAAGAGAG 360  
DB 313 CTGCAACAAGTCAAGAGAGAGCAACAAGTCCAGCAAGAACCCAGAGGCCAAGAGAG 372  
QY 361 GCGGAGCGCAAGGTGAGCCAGAACTACCCCATCTGTGCAAGAACTGTCAGAGGCGCAATGCTG 420  
DB 373 GCGGAGCGCAAGGTGAGCCAGAACTACCCCATCTGTGCAAGAACTGTCAGAGGCGCAATGCTG 432

QY 421 CACGAGCCATCAGCCCGCGCAACCTGAACGCTGGTGAAGGTGATCGAGAGAAAGGCC 480  
DB 433 CACGAGCCCATCAGCCCGCGCAACCTGAATGATGGTGAAGGTGTGAGAGAAAGGCC 492  
QY 481 TTCAGCCCGAGGTGATCCCATATGTTCAACGCGCTGAGCGAGGGCGCAACCCCGAGGAC 540  
DB 493 TTCAGCCCGAGGTGATCCCATATGTTCAACGCGCTGAGCGAGGGCGCAACCCCGAGGAC 552  
QY 541 CTGAACAACAATGTTGAACAACGCTGGCGCGCAACGAGCGCGCATGCAAGATCTGAAGGAC 600  
DB 553 CTGAACAACAATGTTGAACAACGCTGGCGCGCAACGAGCGCGCATGCAAGATCTGAAGGAC 612  
QY 601 ACCATCAACGAGAGGCGCGCGAGTGGAGCGGCTGCAACCCGCTGAGCGCGCGCGCGCTG 660  
DB 613 ACCATCAACGAGAGGCGCGCGAGTGGAGCGGCTGCAACCCGCTGAGCGCGCGCGCGCTG 672  
QY 661 GCGCGCGCGCGAGATGCGCGCAACCCCGCGCGCAACATCGCGCGCGCGCAACCGCTG 720  
DB 673 GCGCGCGCGCGAGATGAGGAGCGCGCGCAACATCGCGCGCGCGCAACCGCTG 732  
QY 721 CAGGAGCAAGTCCCTGATGATGACCAACACCCCGCTGCGCGTGGCGCAACTCAAG 780  
DB 733 CAGGAGCAAGTCCCTGATGATGACCAACACCCCGCTGCGCGTGGCGCAACTCAAG 792  
QY 781 CGGTGATCATCTTGGGCGCTGAACAAGATGCTGCGGATGTACAGCCCGTGAACATCTG 840  
DB 793 AGGTGATCATCTTGGGCGCTGAACAAGATGCTGAGAGATGTACAGCCCGTGAACATCTG 852  
QY 841 GACATCCG 900  
DB 853 GACATCAAGCGAGGCG 912  
QY 901 CTGCGCGCGCGCGAGCGCGCAACCGAGAGAGCTGAAGAACTGATGACCGAGACCTGCTG 960  
DB 913 CTGAGGCGCGAGCG 972  
QY 961 CAGAACCGCAACCCCGAGCTGCAAGAACCATCTGCGCGCGCTCTCGCGCGCGCGCGCGCG 1020  
DB 973 CAGAACCGCAACCCCGAGCTGCAAGAACCATCTGAGGCGCGCGCGCGCGCGCGCGCGCG 1032  
QY 1021 GAGGAGATGATGACCGCGCTGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1080  
DB 1033 GAGGAGATGATGACCGCGCTGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1092  
QY 1081 GCGGAGCGGATGAGCGCGCGCAACGCGTGAACATCATGATGCAAGAGCAACTTCAAG 1140  
DB 1093 GCGGAGCGGATGAGCGCGCGCAACGCGTGAACATCATGATGCAAGAGCAACTTCAAG 1149  
QY 1141 GCGCGCGCGCGCAACGCTCAAGTCTTCAACTGCGCGCAAGAGGCGCGCATCGCGCAAGAC 1200  
DB 1150 GCGAGCAAGAGCATGCTGAAGTCTTCAACTGCGCGCAAGAGGCGCGCATCGCGCAAGAC 1209  
QY 1201 TGCGCGCGCGCGCGCAAGAGGCTGCTGGAATGCGCGCAAGAGGCGCGCATCGATGAAG 1260  
DB 1210 TGCAAGGCGCGCGCGCAAGAGGCTGCTGGAATGCGCGCAAGAGGCGCGCATCGATGAAG 1269  
QY 1261 GACTGCAACGAGCGCGCAAGGCTTCTGCGCGCAAGATCTGCGCGCAAGAGGCGCGC 1320  
DB 1270 GACTGCAACGAGCGCGCAAGGCTTCTGCGCGCAAGATCTGCGCGCAAGAGGCGCGC 1329  
QY 1321 CCGCGCGCACTTCTGCAAGAACCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380  
DB 1330 CCGCGCGCACTTCTGCAAGAACG-----GCCGAGCGCGCGCGCGCGCGCGCGCGCG 1371  
QY 1381 CCGCGCGAGAGCTTCCGCTTGAAGAGAACACCCCGCGCGCGCGCGCGCGCGCGCGCGCG 1440  
DB 1372 CCGGAGGAGAGCTTCAAGTTCGAGGAGAGAGAACCAACCCCGCGCGCGCGCGCGCGCGCG 1431  
QY 1441 CCGGAGCGCTACCGCGAGCG 1500  
DB 1432 ATGCAACAAGAGCTGTACCCCGCTGACAGCGCTGAAGAGCGCTGTCGCGCAACGACCCAGC 1491



QY 1501 AGCCAGTAA 1509  
Db 1492 AGCCAGGAA 1500

RESULT 13

LOCUS AF201927 1509 bp DNA linear SYN 16-MAR-2000  
DEFINITION Synthetic construct gag protein gene, complete cds.

ACCESSION AF201927  
VERSION AF201927.1 GI:7248702

KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
artificial sequences.

REFERENCE 1 (bases 1 to 1509)  
AUTHORS zur Megede,J., Chen,M.C., Doe,B., Schaefer,M., Greer,C.B.,  
Selby,M., Otten,G.R. and Barnett,S.W.  
Increased expression and immunogenicity of sequence-modified human  
immunodeficiency virus type 1 gag gene

JOURNAL J. Virol. 74 (6), 2628-2635 (2000)  
MEDLINE 20148954  
PubMed 10684277

REFERENCE 2 (bases 1 to 1509)  
AUTHORS zur Megede,J. and Barnett,S.W.  
Direct Submission

JOURNAL Submitted (04-NOV-1999) Vaccines, Chiron Corporation, 4560 Horton,  
Emeryville, CA 94608, USA

FEATURES  
Source location/Qualifiers

1..1509  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
1..1509  
/note="HIV-1SP2 p55gag"  
/codon\_start=1  
/transl\_table=11  
/product="gag protein"  
/protein\_id="AAF43628.1"  
/db\_xref="GI:7248703"

/translation="MGARASVSGSLDQKWKIRLRPGKKYKLRHIVWASRELERF  
AVNPGLTSEBGRQILQLOSLQTSBBLSLNVTATLYCHQRIDVKOTYALR  
KIEEONKSKKAQAAAAAGTGNSSOVNYPVQNLQGMVHQAISPRILNWKV  
VBEKAFSPBVIEMFSAISBGAIFQDLNLTMLNTVGHQAMQMLKETINEBAERDVH  
PVHAGPIAPGOMRBPGRSDIAGTSTLQBOIGMTNNPPIPVGEIYKRWIILGLNKIV  
RMVSPSTILDIRGPKBPRFYDRPYKTLRAEQASQDVKNMTETLLVQNAEDCKT  
ILKALGPAATLEMTACOGVGGHKAQVLAEMSOVTNPATIMQORGNFRNDRKT  
KCFNCGKEGHTAHNCRAFRKKGWRCGRGHQMDCTERQANFLKIMPSYKGRPNF  
LQSRBPPTAPRPSFRFGKRTTPSQOKRPIDKELPLTSLRSLFGNDPSQ"

ORIGIN

Query Match 79.6%; Score 1201; DB 12; Length 1509;  
Best Local Similarity 88.5%; Pred. No. 4.1e-118;  
Matches 1351; Conservative 0; Mismatches 140; Indels 36; Gaps 3;

QY 1 ATGGGCGCGCGCGCAGCATCTGCGCGCGAGAGCTGGAAGAAGTGGAGAAGATCCGC 60  
Db 1 ATGGGCGCGCGCGCAGCGCTGCTAGCGCGCGAGCTGGAAGAAGTGGAGAAGATCCGC 60  
QY 61 CTGCGCGCGCGCGCGCAAGAAGCACTACATGCTGAAGCACTGCTGTGGCCAGCCGCGAG 120  
Db 61 CTGCGCGCGCGCGCGCAAGAAGCACTACATGCTGAAGCACTGCTGTGGCCAGCCGCGAG 120  
QY 121 CTGAGGCGCTTCCGCTGTAACCCCGGCTGCTGAGAGACCGCGAGGGCTGCAAGCATC 180  
Db 121 CTGAGGCGCTTCCGCTGTAACCCCGGCTGCTGAGAGACCGCGAGGGCTGCGCCAGATC 180  
QY 181 ATGAAGCACTGCAAGCCCGCTGCAAGCCCGCAACCGAGAGCTGCGCAGCTGTACAAC 240  
Db 181 CTGGGCCAGCTGCAAGCCCGCTGCAAGCCCGCAACCGAGAGCTGCGCAGCTGTACAAC 240  
QY 241 ACCGTGGCAACCTGTACTGCTGTCAGCGCGCATGAGTCCGCGACACCAAGAGGCC 300  
Db 241 ACCGTGGCAACCTGTACTGCTGTCAGCGCGCATGAGTCCGCGACACCAAGAGGCC 300

Db 241 ACCGTGGCAACCTGTACTGCTGTCAGCGCGCATGAGTCCGCGACACCAAGAGGCC 300  
QY 301 CTGGAACAAGATCGAGAGAGACAGAACAGTCCAGCAGAGAACCCAGCGCCAAAGAG 360  
Db 301 CTGGAACAAGATCGAGAGAGACAGAACAGTCCAGAGAACAGAGCCAGCGCCCGCC 360  
QY 361 GCCGACGCA-----AGGTGAGCCAGAACTAACCCTATCGTGCAGAACTCTG 405  
Db 361 GCCGACGCAACCCGCAACAGCAGCCAGGTGAGCCAGAACTAACCCTATCGTGCAGAACTCTG 420  
QY 406 CAGGCGCAGATGTGTCACCAAGCCATCAAGCCCCCGCAACCCCTGAACGCTGTGAAGTG 465  
Db 421 CAGGCGCAGATGTGTCACCAAGCCATCAAGCCCCCGCAACCCCTGAACGCTGTGAAGTG 480  
QY 466 ATGAGAGAGAGGCTTCAAGCCCCCGAGGTGATCCCATGTTCAACCGCCCTGAGCGAGGCG 525  
Db 481 GTGAGAGAGAGGCTTCAAGCCCCCGAGGTGATCCCATGTTCAACCGCCCTGAGCGAGGCG 540  
QY 526 GCCACCCCCCAGAGCTGAAACAAGATGTTGAACACCGTGGCGGCAACAGGCGCCATG 585  
Db 541 GCCACCCCCCAGAGCTGAAACAAGATGTTGAACACCGTGGCGGCAACAGGCGCCATG 600  
QY 586 CAGATGCTGAAGGACCATCAACGAGAGGCGCGGAGTGGGACCGCTTGCAACCCCGTG 645  
Db 601 CAGATGCTGAAGGAGACCATCAACGAGAGGCGCGGAGTGGGACCGCTTGCAACCCCGTG 660  
QY 646 CAGGCGCGCGCGTGGCGCGCGCGCAGATGCGGACCCCGCGGCAAGCATCGCCGCG 705  
Db 661 CAGGCGCGCGCGTGGCGCGCGCGCAGATGCGGACCCCGCGGCAAGCATCGCCGCG 720  
QY 706 GCCACACGACCCCTGCAAGAGCAGATGCTGATGACCAAGCAACCCCGCTGCGCGTG 765  
Db 721 ACCACACGACCCCTGCAAGAGCAGATGCTGATGACCAAGCAACCCCGCTGCGCGTG 780  
QY 766 GCGGACATCTAACAAGCGGTGATCATCTGGGCTTGAACAAGATCTGCGATGTACAGC 825  
Db 781 GCGGACATCTAACAAGCGGTGATCATCTGGGCTTGAACAAGATCTGCGATGTACAGC 840  
QY 826 CCGGTGAGATCTGGAACATCCGCGAGGCGCCCAAGGAGCCTTCCGCACTACGTGAC 885  
Db 841 CCGGTGAGATCTGGAACATCCGCGAGGCGCCCAAGGAGCCTTCCGCACTACGTGAC 900  
QY 886 CGCTTCTTAAGAACCCTGCGCGCGGAGCAGGCGCAACCGAGACGTGAAGAACTGATGAC 945  
Db 901 CGCTTCTTAAGAACCCTGCGCGCGGAGCAGGCGCAACCGAGACGTGAAGAACTGATGAC 960  
QY 946 GAGACCTGCTGTGACAGACGCGCAACCCGACTGCAAGACATCTGCGCTCTGCGC 1005  
Db 961 GAGACCTGCTGTGACAGACGCGCAACCCGACTGCAAGACATCTGCGCTCTGCGC 1020  
QY 1006 CCGGCGCGCAACCTGAGAGAGATGATACCGCTGCAAGGCGTGGCGCGCGCCAC 1065  
Db 1021 CCGGCGCGCAACCTGAGAGAGATGATACCGCTGCAAGGCGTGGCGCGCGCCAC 1080  
QY 1066 AAGGCGCGGTGCTGCGAGGCGATGAGCCAGG---CCAACAGCGTGAACATCATGATG 1122  
Db 1081 AAGGCGCGGTGCTGCGAGGCGATGAGCCAGTGAAGAACCCGCGACCATCATGATG 1140  
QY 1123 CAGAAAGCAACTTCAAGGCGCGCGCGCAAGTCAAGTCTTCAACTGCGCAAGAG 1182  
Db 1141 CAGGCGGCAACTTCCGAACAGCGGAAGACGTCAAGTCTTCAACTGCGCAAGAG 1200  
QY 1183 GCGCACATGCGCAAGAACTGCGCGCGCCCGCAAGAAAGGCTGTAAGTGGCGCAAG 1242  
Db 1201 GCGCACATGCGCAAGAACTGCGCGCGCCCGCAAGAAAGGCTGTAAGTGGCGCGCG 1260  
QY 1243 GAGGCGACAGATGAAGGACTGCAACGAGCGCAGGCGCAACTTCTGCGCAAGATCTGG 1302  
Db 1261 GAGGCGACAGATGAAGGACTGCAACGAGCGCAGGCGCAACTTCTGCGCAAGATCTGG 1320  
QY 1303 CCCAGCCAAAGGCGCGCGCGCAACTTCTGCAAGAACCGCAAGCGAGCGCGCGCGCC 1362  
Db 1321 CCCAGCCAAAGGCGCGCGCGCAACTTCTGCAAGAGCG-----C 1362







Wed Jun 2 09:56:09 2004

us-09-475-704a-4.rge

Page 16

Search completed: June 1, 2004, 13:09:33  
Job time : 5857 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: May 28, 2004, 11:33:50 ; Search time 758.462 Seconds  
(without alignments)  
6870.358 Million cell updates/sec

Title: US-09-475-704A-3  
Perfect score: 1479  
Sequence: 1 atgggccccgcgcgcagcat.....acgacccctgagccagtaaa 1479

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2960401 seqs, 2274450654 residues

Total number of hits satisfying chosen parameters: 5920802

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database :

Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 18: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 19: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1479	100.0	1479	10	US-09-967-464-63 Sequence 63, Appl
2	1479	100.0	1479	13	US-09-899-575-3 Sequence 3, Appl
3	1477.4	99.9	3162	15	US-10-190-435-18 Sequence 18, Appl
4	1477.4	99.9	3462	15	US-10-190-435-16 Sequence 16, Appl
5	1477.4	99.9	4419	15	US-10-190-435-19 Sequence 19, Appl
6	1477.4	99.9	4419	15	US-10-190-305A-14 Sequence 14, Appl
7	1477.4	99.9	4483	15	US-10-190-435-35 Sequence 35, Appl
8	1477.4	99.9	4606	15	US-10-190-435-34 Sequence 34, Appl
9	1477.4	99.9	4615	15	US-10-190-435-36 Sequence 36, Appl
10	1477.4	99.9	4702	15	US-10-190-435-38 Sequence 38, Appl
11	1477.4	99.9	4716	15	US-10-190-435-17 Sequence 17, Appl
12	1477.4	99.9	4716	15	US-10-190-305A-13 Sequence 13, Appl
13	1475.8	99.8	2742	15	US-10-190-435-20 Sequence 20, Appl
14	1475.8	99.8	2742	15	US-10-190-305A-15 Sequence 15, Appl

	15	1475.8	99.8	3930	15	US-10-190-435-9	Sequence 9, Appl
	16	1475.8	99.8	3930	15	US-10-190-435-10	Sequence 10, Appl
	17	1475.8	99.8	3930	15	US-10-190-435-11	Sequence 11, Appl
	18	1475.8	99.8	5145	15	US-10-190-435-12	Sequence 12, Appl
	19	1475.8	99.8	5145	15	US-10-190-305A-12	Sequence 12, Appl
	20	1474.8	99.7	4713	15	US-10-190-435-59	Sequence 59, Appl
	21	1474.8	99.7	4713	15	US-10-190-305A-83	Sequence 83, Appl
	22	1473.8	99.6	5184	15	US-10-190-435-58	Sequence 58, Appl
	23	1473.8	99.6	5184	15	US-10-190-305A-82	Sequence 82, Appl
	24	1473.4	99.6	1479	10	US-09-967-464-67	Sequence 67, Appl
	25	1472.4	99.6	2742	15	US-10-190-435-57	Sequence 57, Appl
	26	1472.4	99.6	2742	15	US-10-190-305A-81	Sequence 81, Appl
	27	1463	98.9	1479	13	US-09-899-575-20	Sequence 20, Appl
	28	1315.2	88.9	3531	15	US-10-190-435-13	Sequence 20, Appl
	29	1315.2	88.9	3537	15	US-10-190-435-14	Sequence 14, Appl
	30	1315.2	88.9	3537	15	US-10-190-435-15	Sequence 15, Appl
	31	1288.8	87.1	1491	13	US-09-899-575-99	Sequence 99, Appl
	32	1288	87.1	1494	13	US-09-899-575-51	Sequence 51, Appl
	33	1276.8	86.3	1509	13	US-09-899-575-4	Sequence 4, Appl
	34	1275.2	86.2	1509	10	US-09-967-464-64	Sequence 64, Appl
	35	1273.6	86.1	1509	10	US-09-967-464-68	Sequence 68, Appl
	36	1260.8	85.2	1509	13	US-09-899-575-21	Sequence 21, Appl
	37	1221.2	82.6	9166	13	US-10-359-120-168	Sequence 168, App
	38	1206	81.5	9788	13	US-10-359-120-174	Sequence 174, App
	39	1197.8	81.0	1515	13	US-10-387-336-4	Sequence 4, Appl
	40	1197.8	81.0	4472	13	US-10-387-336-75	Sequence 75, Appl
	41	1197.8	81.0	4608	13	US-10-387-336-76	Sequence 76, Appl
	42	1197.8	81.0	4689	13	US-10-387-336-74	Sequence 74, Appl
	43	1197.8	81.0	4766	13	US-10-387-336-73	Sequence 73, Appl
	44	1196.2	80.9	2799	15	US-10-241-009-18	Sequence 18, Appl
	45	1196.2	80.9	2799	15	US-10-190-434B-18	Sequence 18, Appl

## ALIGNMENTS

RESULT 1  
US-09-967-464-63  
; Sequence 63, Application US/09967464  
; Publication No. US20030138453A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Hagan, Derek  
; APPLICANT: Otten, Gillis  
; APPLICANT: Donnelly, John J.  
; APPLICANT: Barnett, Susan  
; APPLICANT: Singh, Mamohan  
; APPLICANT: Ulmer, Jeffrey  
; APPLICANT: Dubensky, Jr., Thomas W.  
; TITLE OF INVENTION: MICROPARTICLES FOR DELIVERY OF HETEROLOGOUS NUCLEIC ACIDS  
; FILE REFERENCE: PPI6269.004  
; CURRENT APPLICATION NUMBER: US/09/967,464  
; PRIOR FILING DATE: 2002-04-11  
; PRIOR APPLICATION NUMBER: 60/236,105  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: 60/315,905  
; PRIOR FILING DATE: 2001-08-30  
; NUMBER OF SEQ ID NOS: 68  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 63  
; LENGTH: 1479  
; TYPE: DNA  
; ORGANISM: Human immunodeficiency virus type 1  
US-09-967-464-63

Query Match 100.0%; Score 1479; DB 10; Length 1479;  
Best Local Similarity 100.0%; Pred. No. 6.9e-303;  
Matches 1479; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGGCGCGCGCGCAGCATCTGCGCGCGCGCAAGCTGAGACGCTGGAGCGCATCCGC 60  
Db 1 ATGGGCGCGCGCGCAGCATCTGCGCGCGCGCAAGCTGAGACGCTGGAGCGCATCCGC 60

```
QY 61 CTGCGCCCGCGCGAAGATGCTACATGATGAAGCACTGTGTGGCCAGCCCGGAG 120
DB 61 CTGCGCCCGCGCGCAAGATGCTACATGATGAAGCACTGTGTGGCCAGCCCGGAG 120
QY 121 CTGGAAGATTGCGCCCTGAACCCCGGCTCTGTGAGACAGAGGGCTGCAAGCATC 180
DB 121 CTGGAAGATTGCGCCCTGAACCCCGGCTCTGTGAGACAGAGGGCTGCAAGCATC 180
QY 181 ATCCGCAAGCTGCAACCCCGCTGCAAGACCGGCAAGAGAGCTGAAGAGCTGTCAAC 240
DB 181 ATCCGCAAGCTGCAACCCCGCTGCAAGACCGGCAAGAGAGCTGAAGAGCTGTCAAC 240
QY 241 ACCGTGGCCACCTGTACTGTGCTGCAAGAGATCGAGGTCCGCAACCAAGAGGCC 300
DB 241 ACCGTGGCCACCTGTACTGTGCTGCAAGAGATCGAGGTCCGCAACCAAGAGGCC 300
QY 301 CTGGAACAAGATCGAGAGAGCAAGAACTGCTGCAAGAGATCCAGAGCCGAGGCC 360
DB 301 CTGGAACAAGATCGAGAGAGCAAGAACTGCTGCAAGAGATCCAGAGCCGAGGCC 360
QY 361 GCCGACAAGGGCAAGGTGAGCGCAAGAACTACCTCATCTGTGCAAGACTGTGAGG 420
DB 361 GCCGACAAGGGCAAGGTGAGCGCAAGAACTACCTCATCTGTGCAAGACTGTGAGG 420
QY 421 GTGCAACCAAGCATCAGCCCGCGCAACCTGTAACCTGTGTGAAGGTATCGAGAGAG 480
DB 421 GTGCAACCAAGCATCAGCCCGCGCAACCTGTAACCTGTGTGAAGGTATCGAGAGAG 480
QY 481 GCCTTCAGCCCGGAGGTATCCCATATGTTCAACCGCCCTGAGGAGGGCGCCACCC 540
DB 481 GCCTTCAGCCCGGAGGTATCCCATATGTTCAACCGCCCTGAGGAGGGCGCCACCC 540
QY 541 GACCTGAACAAGATTGTAACAACCTGTGGCGGCAACAGGCCCATGCAATGCTGAG 600
DB 541 GACCTGAACAAGATTGTAACAACCTGTGGCGGCAACAGGCCCATGCAATGCTGAG 600
QY 601 GACACCATCAACGAGAGAGCGCGGAGTGGAACCGGCGTGCACCCCGTGCACCGCC 660
DB 601 GACACCATCAACGAGAGAGCGCGGAGTGGAACCGGCGTGCACCCCGTGCACCGCC 660
QY 661 ATGCCCCCGGCGAGATGCGGCAAGCCCGGCGAGCATGCGCGCACCAAGCACCC 720
DB 661 ATGCCCCCGGCGAGATGCGGCAAGCCCGGCGAGCATGCGCGCACCAAGCACCC 720
QY 721 CTGCAAGAGAGATGCGCTGTGATGACCAACCCCGCATCCCGTGGCGCATCTAC 780
DB 721 CTGCAAGAGAGATGCGCTGTGATGACCAACCCCGCATCCCGTGGCGCATCTAC 780
QY 781 AAGCGGTGATCATCTGGGCTGAAACAAGATCGTGCGATGTAACGCCCGTGAATC 840
DB 781 AAGCGGTGATCATCTGGGCTGAAACAAGATCGTGCGATGTAACGCCCGTGAATC 840
QY 841 CTGCAATCAAGCAGGCGCCCAAGAGCCCTTCGCGACTAGCTGAGCCGCTTCAAG 900
DB 841 CTGCAATCAAGCAGGCGCCCAAGAGCCCTTCGCGACTAGCTGAGCCGCTTCAAG 900
QY 901 ACCCTGCGCGCGAGCAGAGCAACCAAGAGTGAAGAACTGATGACCGACCTGCTG 960
DB 901 ACCCTGCGCGCGAGCAGAGCAACCAAGAGTGAAGAACTGATGACCGACCTGCTG 960
QY 961 GTGCAAGAACCGCAACCCCGCTGCAAGAACCTGCGCGCTCTCGGCCCGCGCGCAG 1020
DB 961 GTGCAAGAACCGCAACCCCGCTGCAAGAACCTGCGCGCTCTCGGCCCGCGCGCAG 1020
QY 1021 CTGGAAGAGTATGACCGCTGTGCAAGAGCGGCGCGCGCGCGCGCGCGCGCTG 1080
DB 1021 CTGGAAGAGTATGACCGCTGTGCAAGAGCGGCGCGCGCGCGCGCGCGCGCTG 1080
QY 1081 CTGCGCGAGGCGATGAGCGCAACCAAGCGTATGATGCAAGAGCACTTCAAG 1140
DB 1081 CTGCGCGAGGCGATGAGCGCAACCAAGCGTATGATGCAAGAGCACTTCAAG 1140
QY 1141 GGGCCCCGGCGATGCTCAAGTCTTCACTGGCGCAAGAGGCCCATCGCCCGCAAC 1200
```

```
DB 1141 GGGCCCCGGCGATGCTCAAGTCTTCACTGGCGCAAGAGGCCCATCGCCCGCAAC 1200
QY 1201 TGCCGCGCCCCCGCAAGAGGGCTGTGAAAGTGCAGAGAGGGCCACCATGTAAG 1260
DB 1201 TGCCGCGCCCCCGCAAGAGGGCTGTGAAAGTGCAGAGAGGGCCACCATGTAAG 1260
QY 1261 GACTGCACGAGCGCGCAAGCTTCTGTGGCAAGATCTGGCCAGCCCAAGGGCCGC 1320
DB 1261 GACTGCACGAGCGCGCAAGCTTCTGTGGCAAGATCTGGCCAGCCCAAGGGCCGC 1320
QY 1321 CCCGCAACTTCTGCAAGACCGCCCGAGCCCAACCGCCCGCCCGCGAGCTTCCGC 1380
DB 1321 CCCGCAACTTCTGCAAGACCGCCCGAGCCCAACCGCCCGCCCGCGAGCTTCCGC 1380
QY 1381 TTGAGAGAGCAACCCCGCGCAAGAGAGAGAGCAAGAGCCGAGACCTGACAGC 1440
DB 1381 TTGAGAGAGCAACCCCGCGCAAGAGAGAGAGCAAGAGCCGAGACCTGACAGC 1440
QY 1441 CTGAAGAGCTGTTCGCAAGACCCCTGAGCCAGTAA 1479
DB 1441 CTGAAGAGCTGTTCGCAAGACCCCTGAGCCAGTAA 1479
```

## RESULT 2

US-09-899-575-3

; Sequence 3, Application US/09899575

; Publication No. US20030223961A1

; GENERAL INFORMATION:

; APPLICANT: Zur Megele, Jan

; APPLICANT: Barnett, Susan W.

; APPLICANT: Egnelbrecht, Susan

; APPLICANT: van Rensburg, Estrelita Janse

; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C

; FILE REFERENCE: PP01631.102

; CURRENT APPLICATION NUMBER: US/09/899,575

; PRIOR FILING DATE: 2001-07-05

; PRIOR APPLICATION NUMBER: 09/475,704

; NUMBER OF SEQ ID NOS: 135

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 3

; LENGTH: 1479

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: synthetic Gag

; US-09-899-575-3

Query Match 100.0%; Score 1479; DB 13; Length 1479;

Best Local Similarity 100.0%; Pred. No. 6.9e-303;

Matches 1479; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGGGCGCCCGCGCAGCATCTGCGCGCGCGCAAGCTGAGCGCTGGAGCGATCCGC 60
DB 1 ATGGGCGCCCGCGCAGCATCTGCGCGCGCGCAAGCTGAGCGCTGGAGCGATCCGC 60
QY 61 CTGCGCCCGCGCGCAAGAGTGTACTATGATGAAGCACTGTGTGGCGCGCGGAG 120
DB 61 CTGCGCCCGCGCGCAAGAGTGTACTATGATGAAGCACTGTGTGGCGCGCGGAG 120
QY 121 CTGGAAGAGTTCGCTGTAACCCCGCGCTGTGAGAGCAAGAGGGCTGCAAGCATC 180
DB 121 CTGGAAGAGTTCGCTGTAACCCCGCGCTGTGAGAGCAAGAGGGCTGCAAGCATC 180
QY 181 ATCCGCAAGCTGCAACCCCGCTGCAAGACCGGCAAGAGAGCTGAAGAGCTGTCAAC 240
DB 181 ATCCGCAAGCTGCAACCCCGCTGCAAGACCGGCAAGAGAGCTGAAGAGCTGTCAAC 240
QY 241 ACCGTGGCCACCTGTACTGTGCTGCAAGAGATCGAGGTCCGCAACCAAGAGGCC 300
```



```
Db 241 ACCGTGGCCACCCTGTACTGCTGTGACGAGAAAGATCGAGTCCGCCACCAAGAGGCC 300
QY 301 CTGGAAGATCGAGAGAGAGAGAGAGAGAGATCCAGCAGAGGCCGAGGCC 360
Db 301 CTGGAAGATCGAGAGAGAGAGAGAGAGAGATCCAGCAGAGGCCGAGGCC 360
QY 361 GCCGCAAGGGGCAAGGTGAGCCAGAACTAATCCATCTGTGAGAACTGTGAGGGCCAGATG 420
Db 361 GCCGCAAGGGGCAAGGTGAGCCAGAACTAATCCATCTGTGAGAACTGTGAGGGCCAGATG 420
QY 421 GTGACCAAGGCCATCAGCCCCCGCACCTGAAACGCTGTGGTGAAGGTGATCGAGAGAA 480
Db 421 GTGACCAAGGCCATCAGCCCCCGCACCTGAAACGCTGTGGTGAAGGTGATCGAGAGAA 480
QY 481 GCCTTCAGCCCCGAGGTGATCCCATGTTCACCGCCCTGAGCGAGGGCCACCCCCCAG 540
Db 481 GCCTTCAGCCCCGAGGTGATCCCATGTTCACCGCCCTGAGCGAGGGCCACCCCCCAG 540
QY 541 GACCTGAACAGATGTGAACAACGTGGCGGCCACCAAGGCCGATGAGATGCTGAAG 600
Db 541 GACCTGAACAGATGTGAACAACGTGGCGGCCACCAAGGCCGATGAGATGCTGAAG 600
QY 601 GACACCATCAACGAGAGGGCGCGCGAGGTGGAGCCGCTGCAACCCCGTGAACGCGCGCCC 660
Db 601 GACACCATCAACGAGAGGGCGCGCGAGGTGGAGCCGCTGCAACCCCGTGAACGCGCGCCC 660
QY 661 ATGCCCCCGGCGCAGATGCGCGAGAGCCCGCGCAGCGACATCGCCGCGCACCAAGAC 720
Db 661 ATGCCCCCGGCGCAGATGCGCGAGAGCCCGCGCAGCGACATCGCCGCGCACCAAGAC 720
QY 721 CTGCAAGAGAGATCGCTGTGATGACGAGCAACCCCGCATCCCGTGGCGGACATCTAC 780
Db 721 CTGCAAGAGAGATCGCTGTGATGACGAGCAACCCCGCATCCCGTGGCGGACATCTAC 780
QY 781 AAGCGGTGATCATCTGGGGCTGAAACAAGATCGTGGATGTACAGCCCCGTGAGATC 840
Db 781 AAGCGGTGATCATCTGGGGCTGAAACAAGATCGTGGATGTACAGCCCCGTGAGATC 840
QY 841 CTGGAATCAAGAGAGGCCCCCAAGAGCCCTTCCGCACTACGTGACCTTCTTCAAG 900
Db 841 CTGGAATCAAGAGAGGCCCCCAAGAGCCCTTCCGCACTACGTGACCTTCTTCAAG 900
QY 901 ACCCTGCGCGCGGAGAGAGCAACCCAGAGGTGAAGAACTGGATGACCAACCTCTG 960
Db 901 ACCCTGCGCGCGGAGAGAGCAACCCAGAGGTGAAGAACTGGATGACCAACCTCTG 960
QY 961 GTGCAAGAACGCCAACCCCGACTGCAAGACCATCTGCGCGCTCTCGGCCCGCGCCAGC 1020
Db 961 GTGCAAGAACGCCAACCCCGACTGCAAGACCATCTGCGCGCTCTCGGCCCGCGCCAGC 1020
QY 1021 CTGGAAGAGATGATACCGCTGCGCAGAGCGGTGGCGGCCCGCAGCCACAAGGCCGCTG 1080
Db 1021 CTGGAAGAGATGATACCGCTGCGCAGAGCGGTGGCGGCCCGCAGCCACAAGGCCGCTG 1080
QY 1081 CTGCGCGAGGCGGATGAGCCAGGCCACAACAGCGTGAATGATGCAAGAAAGCACTTCAAG 1140
Db 1081 CTGCGCGAGGCGGATGAGCCAGGCCACAACAGCGTGAATGATGCAAGAAAGCACTTCAAG 1140
QY 1141 GGGCCCCCGCGCATGTCTCAAGTGTCTTCACTGCGGCAAGAGGGCCACATGCGCCGCAAC 1200
Db 1141 GGGCCCCCGCGCATGTCTCAAGTGTCTTCACTGCGGCAAGAGGGCCACATGCGCCGCAAC 1200
QY 1201 TGGCGCGCGCGCGCGAGAAAGGCTGTGAAGTGGCGCAAGAGGGCCACCAAGATGAAG 1260
Db 1201 TGGCGCGCGCGCGCGAGAAAGGCTGTGAAGTGGCGCAAGAGGGCCACCAAGATGAAG 1260
QY 1261 GACTGCAACGAGCGCCAGGCCAATCTTCTTGGGCAAGATCTGGCCAGCCACAAGGGCGCG 1320
Db 1261 GACTGCAACGAGCGCCAGGCCAATCTTCTTGGGCAAGATCTGGCCAGCCACAAGGGCGCG 1320
QY 1321 CCGGCAACTTCTTCTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
Db 1321 CCGGCAACTTCTTCTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
```

```
QY 1381 TTGAGAGACACACCCCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
Db 1381 TTGAGAGAGACACACCCCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
QY 1441 CTGAAGAGCTGTTGCGCAACGAGACCCCTGAGCCAGTAA 1479
Db 1441 CTGAAGAGCTGTTGCGCAACGAGACCCCTGAGCCAGTAA 1479

RESULT 3
US-10-190-435-18
; Sequence 18, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGBERCHT, Susan
; APPLICANT: VAN RENSBURG, Betrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: P18133.003 / 2302-18133
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 3162
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GagRTmut_C
US-10-190-435-18

Query Match 99.9%; Score 1477.4; DB 15; Length 3162;
Best Local Similarity 99.9%; Pred. No. 1.5e-302;
Matches 1478; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGCGCGCGCAGCATCTGCGCGCGCGCAAGCTGGAACGCTGGAGCGCATCCGC 60
Db 7 ATGGGCGCGCGCGCAGCATCTGCGCGCGCGCAAGCTGGAACGCTGGAGCGCATCCGC 66
QY 61 CTGCGCGCGCGCGCGCAAGAGTGTACATGATGAGCACTGTGTGGGCCAGCGCGAG 120
Db 67 CTGCGCGCGCGCGCGCAAGAGTGTACATGATGAGCACTGTGTGGGCCAGCGCGAG 126
QY 121 CTGGAAGATTGCGCTGAACCCCGGCTGCTGTGAGACCAAGAGGCTGCAAGCAGATC 180
Db 127 CTGGAAGATTGCGCTGAACCCCGGCTGCTGTGAGACCAAGAGGCTGCAAGCAGATC 186
QY 181 ATCCGCAAGCTGCAACCCCGCTGCAAGCCGCGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
Db 187 ATCCGCAAGCTGCAACCCCGCTGCAAGCCGCGAGAGAGAGAGAGAGAGAGAGAGAGAG 246
QY 241 ACCGTGGCACCCCTGTACTGCTGCAAGAGAGATGAGGTCCGCGACCAAGAGAGGCC 300
Db 247 ACCGTGGCACCCCTGTACTGCTGCAAGAGAGATGAGGTCCGCGACCAAGAGAGGCC 306
QY 301 CTGCAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
Db 307 CTGCAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 366
QY 361 GCCGCAAGGGGCAAGGTGAGCCAGAACTAATCCATCTGTGAGAACTGTGAGGGCCAGATG 420
Db 367 GCCGCAAGGGGCAAGGTGAGCCAGAACTAATCCATCTGTGAGAACTGTGAGGGCCAGATG 426
QY 421 GTGCAACAGGCGCATGAGCCCGCGACCTTGAAGCGCTGGGTGAAGGTGATCGAGAGAGAG 480
Db 427 GTGCAACAGGCGCATGAGCCCGCGACCTTGAAGCGCTGGGTGAAGGTGATCGAGAGAGAG 486
QY 481 GCCTTCAGCCCCGAGGTGATCCCAATGTTCAACCGCCTGAGCGAGGGCGCACCCCGCAG 540
Db 481 GCCTTCAGCCCCGAGGTGATCCCAATGTTCAACCGCCTGAGCGAGGGCGCACCCCGCAG 540
```

```
Db 487 GCCTTCAGACCCCGAGGTGATCCCATGTTTCAACCGCCCTGAGCGAGGCGCCACCCCCAG 546
QY 541 GACCTGAACAGATGTTGAACACCGTGGGCGCCACACGAGCCGCGCATGAGTCTGAAG 600
Db 547 GACCTGAACAGATGTTGAACACCGTGGGCGCCACACGAGCCGCGCATGAGTCTGAAG 606
QY 601 GACACCATCAACGAGAGGCGCGGAGTGGGACCGCGTGAACCCCGTGAACCGCGGCCCC 660
Db 607 GACACCATCAACGAGAGGCGCGGAGTGGGACCGCGTGAACCCCGTGAACCGCGGCCCC 666
QY 661 ATGCCCCCGCGCGAGATGCGCGGAGCGCCCGCGGAGCGACATCGCCGCGACACGACACC 720
Db 667 ATGCCCCCGCGCGAGATGCGCGGAGCGCCCGCGGAGCGACATCGCCGCGACACGACACC 726
QY 721 CTGCAAGAGCAGATCGCGCTGATGACCAAGACCCCGCATCCCGTGGCGCATCTAC 780
Db 727 CTGCAAGAGCAGATCGCGCTGATGACCAAGACCCCGCATCCCGTGGCGCATCTAC 786
QY 781 AAGCGGTGATCATCTGCGCTGAAACAAGATCGTGGGATGTACAGCCCGTGAACATC 840
Db 787 AAGCGGTGATCATCTGCGCTGAAACAAGATCGTGGGATGTACAGCCCGTGAACATC 846
QY 841 CTGGAATATCAAGCAGAGGCGCCCAAGAGCCCTTCCGCACTACGTGACCGCTTCTCAAG 900
Db 847 CTGGAATATCAAGCAGAGGCGCCCAAGAGCCCTTCCGCACTACGTGACCGCTTCTCAAG 906
QY 901 ACCCTGCGCGCGAGCAGACACCCAGAGGTGAAGAACTGATGACCGACACCTGCTG 960
Db 907 ACCCTGCGCGCGAGCAGACACCCAGAGGTGAAGAACTGATGACCGACACCTGCTG 966
QY 961 GTGCAAGACCCCAACCCCGCATGCAAGACCATCTGCGCTCTCGCCCGCGCGCAGC 1020
Db 967 GTGCAAGACCCCAACCCCGCATGCAAGACCATCTGCGCTCTCGCCCGCGCGCAGC 1026
QY 1021 CTGGAAGAGATGATGACCGCTGCAAGGCGTGGGCGCGCCCGACCAAGCGCGCGT 1080
Db 1027 CTGGAAGAGATGATGACCGCTGCAAGGCGTGGGCGCGCCCGACCAAGCGCGCGT 1086
QY 1081 CTGCGCGAGCGGATGAGCCAGGCGCAACACCAAGCGTGAATGCAAGAGCACTTCAAG 1140
Db 1087 CTGCGCGAGCGGATGAGCCAGGCGCAACACCAAGCGTGAATGCAAGAGCACTTCAAG 1146
QY 1141 GGGCCCCCGCGCATCGTCAAGTCTTCAACTGCGGCAAGAGGCGCCACATCGCCCGCAAC 1200
Db 1147 GGGCCCCCGCGCATCGTCAAGTCTTCAACTGCGGCAAGAGGCGCCACATCGCCCGCAAC 1206
QY 1201 TGCGCGCGCGCGCGCAAGAGGCGTGTGAAGTGCAGCAAGAGGCGCCACATGATGAAG 1260
Db 1207 TGCGCGCGCGCGCGCAAGAGGCGTGTGAAGTGCAGCAAGAGGCGCCACATGATGAAG 1266
QY 1261 GACTGCAACGAGCGCGCAAGGCGCACTTCTGGGCAAGATCTGCGCCAGCGCAAGGGCGCG 1320
Db 1267 GACTGCAACGAGCGCGCAAGGCGCACTTCTGGGCAAGATCTGCGCCAGCGCAAGGGCGCG 1326
QY 1321 CCGCGCAACTTCTCTGCAAGGCGCGCGCGAGCGCCACCGCCCCCGCGAGAGCTTCCGC 1380
Db 1327 CCGCGCAACTTCTCTGCAAGGCGCGCGCGAGCGCCACCGCCCCCGCGAGAGCTTCCGC 1386
QY 1381 TTGAGAGAGACCAACCCCGCGCGCAAGAGCAGAGCAGCAAGAGACCGCGAGACCTTGAACAGC 1440
Db 1387 TTGAGAGAGACCAACCCCGCGCGCAAGAGCAGAGCAGCAAGAGACCTTGAACAGC 1446
QY 1441 CTGAAGAGCGCTGTTGCGCAAGACCCCGTGAAGCACTAA 1479
Db 1447 CTGAAGAGCGCTGTTGCGCAAGACCCCGTGAAGCACTAA 1485
```

```
RESULT 4
US-10-190-435-16
; Sequence 16, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGBDE, Jan
```

```
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Betrelicia J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: P18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 3462
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: Description of Artificial Sequence: GagProteinRtmult_C
US-10-190-435-16
```

```
Query Match 99.9%; Score 1477.4; DB 15; Length 3462;
Best Local Similarity 99.9%; Pred. No. 1.5e-302;
Matches 1478; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 ATGGGCGCGCGCGCGCAGCATCTGCGCGCGCGCGCAGCTGACCGCTGGAGCGCATCCGC 60
Db 7 ATGGGCGCGCGCGCGCAGCATCTGCGCGCGCGCGCAGCTGACCGCTGGAGCGCATCCGC 66
QY 61 CTGCGCGCGCGCGCGCAGAAAGTGTACTACATGATGAAGCACTGTGTGTGGCCAGCGCGAG 120
Db 67 CTGCGCGCGCGCGCGCAGAAAGTGTACTACATGATGAAGCACTGTGTGTGGCCAGCGCGAG 126
QY 121 CTGGAAGAGTTCGCTTGAACCCCGCGCTGCTGGAAGACCAAGCGAGGCTGCAAGCAGATC 180
Db 127 CTGGAAGAGTTCGCTTGAACCCCGCGCTGCTGGAAGACCAAGCGAGGCTGCAAGCAGATC 186
QY 181 ATCGCGCAGCTGCAACCCCGCGCTGCAAGACCGCGCAGCGAGAGCTGAAGCTGTTCAC 240
Db 187 ATCGCGCAGCTGCAACCCCGCGCTGCAAGACCGCGCAGCGAGAGCTGAAGCTGTTCAC 246
QY 241 ACCGTGGCCACCTGTACTGCTGTCACGAGAAAGATCGAGTCCGCGACACCAAGAGGCC 300
Db 247 ACCGTGGCCACCTGTACTGCTGTCACGAGAAAGATCGAGTCCGCGACACCAAGAGGCC 306
QY 301 CTGGAACAAGATGAGGAGGAGCAGAAACAAGTGCACAGCAAGATCCAGCGCGCGAGGCC 360
Db 307 CTGGAACAAGATGAGGAGGAGCAGAAACAAGTGCACAGCAAGATCCAGCGCGCGAGGCC 366
QY 361 GCGGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420
Db 367 GCGGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 426
QY 421 GTGCAACAGGCGCATCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
Db 427 GTGCAACAGGCGCATCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 486
QY 481 GCCTTCAGCGCGCGAGTGAATCCCATGTTCAACCGCGCTGAGCGAGGCGCGCGCGCGCG 540
Db 487 GCCTTCAGCGCGCGAGTGAATCCCATGTTCAACCGCGCTGAGCGAGGCGCGCGCGCGCG 546
QY 541 GACCTGAACAGATGTTGAACAACGTTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
Db 547 GACCTGAACAGATGTTGAACAACGTTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 606
QY 601 GACACCATCAACGAGAGGCGCGCGCGAGTGGGACCGCGTGAACCCCGTGAACCGCGGCCCC 660
Db 607 GACACCATCAACGAGAGGCGCGCGCGAGTGGGACCGCGTGAACCCCGTGAACCGCGGCCCC 666
QY 661 ATGCCCCCGCGCGAGATGCGCGGAGCGCCCGCGGAGCGACATCGCGCGCGCGCGCGCG 720
Db 667 ATGCCCCCGCGCGAGATGCGCGGAGCGCCCGCGGAGCGACATCGCGCGCGCGCGCGCG 726
QY 721 CTGCAAGAGCAGATCGCGCTGATGACCAAGACCCCGCATCCCGTGGCGCATCTAC 780
```

Db 727 CTGACAGACAGATCGCCTGTGATGACCAAGCAACCCCCCATCCCGTGGCGCATCTAC 786  
QY 781 AAGCGGTGATCATCTGTGGCTTGAACAAGATCTGCGATGTACAGCCCGGTGAGCATC 840  
Db 787 AAGCGGTGATCATCTGTGGCTTGAACAAGATCTGCGATGTACAGCCCGGTGAGCATC 846  
QY 841 CTGACATCAAGCAGGGGCCCAAGAGCCCTTCCGACTACGTGAGCCGCTTCTTCAAG 900  
Db 847 CTGACATCAAGCAGGGGCCCAAGAGCCCTTCCGACTACGTGAGCCGCTTCTTCAAG 906  
QY 901 ACCCTGCGCGCCGAGCAGAGCAACCCAGAGGTGAAGAACTGTGATGACCCGACACCTGCTG 960  
Db 907 ACCCTGCGCGCCGAGCAGAGCAACCCAGAGGTGAAGAACTGTGATGACCCGACACCTGCTG 966  
QY 961 GTGACAGACGCCAACCCCGACTGCAAGACCATCTGCGCGCTCTGCGCCCGCGCGCAGC 1020  
Db 967 GTGACAGACGCCAACCCCGACTGCAAGACCATCTGCGCGCTCTGCGCCCGCGCGCAGC 1026  
QY 1021 CTGAGGAGATGATGACCGCTTGCAGAGGGGTGGCGCGCCGAGCCAGAGCCCGCGCTG 1080  
Db 1027 CTGAGGAGATGATGACCGCTTGCAGAGGGGTGGCGCGCCGAGCCAGAGCCCGCGCTG 1086  
QY 1081 CTGCGCCGAGCGGATGAGCCAGGCCAACACACAGCGGTGATGATGACAGAGCAACTTCAAG 1140  
Db 1087 CTGCGCCGAGCGGATGAGCCAGGCCAACACACAGCGGTGATGATGACAGAGCAACTTCAAG 1146  
QY 1141 GGGCCCCCGCGCATCTGTCAAGTGTCTTCAACTGCGCGAGAGAGGCCCAATGCGCCGCAAC 1200  
Db 1147 GGGCCCCCGCGCATCTGTCAAGTGTCTTCAACTGCGCGAGAGAGGCCCAATGCGCCGCAAC 1206  
QY 1201 TGCCGCGCGCCCCCGCAAGAGGGCTGTGAAAGTGCGCGCAAGAGGGCCACAGATGAAG 1260  
Db 1207 TGCCGCGCGCCCCCGCAAGAGGGCTGTGAAAGTGCGCGCAAGAGGGCCACAGATGAAG 1266  
QY 1261 GACTGCACCGAGCGCCAGGCCAACTTCTGTGGCAAGATCTGGCCAGGCCAAGAGGGCCGC 1320  
Db 1267 GACTGCACCGAGCGCCAGGCCAACTTCTGTGGCAAGATCTGGCCAGGCCAAGAGGGCCGC 1326  
QY 1321 CCGCGCACTTCTGTGACAGCGCCCGCCAGCCACCGCCCGCCCGCGAGAGCTTCCGC 1380  
Db 1327 CCGCGCACTTCTGTGACAGCGCCCGCCAGCCACCGCCCGCCCGCGAGAGCTTCCGC 1386  
QY 1381 TTCGAGAGACCAACCCCGCCAGAGCAGAGAGCAAGGACCGGAGACCTGTGACAGC 1440  
Db 1387 TTCGAGAGACCAACCCCGCCAGAGCAGAGAGCAAGGACCGGAGACCTGTGACAGC 1446  
QY 1441 CTGAAGAGCTGTTCGGCAAGACCCCTGAGCCAGTAA 1479  
Db 1447 CTGAAGAGCTGTTCGGCAAGACCCCTGAGCCAGAAA 1485

RESULT 5  
US-10-190-435-19  
; Sequence 19, Application US/10190435  
; Publication No. US20030143248A1  
; GENERAL INFORMATION:  
; APPLICANT: ZUR MEGEDH, Jan  
; APPLICANT: BARNETT, Susan W.  
; APPLICANT: LIAN, Ying  
; APPLICANT: ENGBRECHT, Susan  
; APPLICANT: VAN RENSBURG, Bastelita J.  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C  
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF  
; FILE REFERENCE: P18133.003 / 2302-18133  
; CURRENT FILING DATE: 2002-12-30  
; NUMBER OF SEQ ID NOS: 319  
; SOFTWARE: Patent Ver. 2.0  
; SEQ ID NO 19  
; LENGTH: 4419  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: GagRtmultatRevnef\_C  
US-10-190-435-19  
Query Match 99.9%; Score 1477.4; DB 15; Length 4419;  
Best Local Similarity 99.9%; Pred. No. 1.4e-302;  
Matches 1478; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 ATGGCGCCCGCGCAGATCTGTGCGCGCGCAAGCTGACGCTGGAGCGCATCCGC 60  
Db 7 ATGGCGCCCGCGCAGATCTGTGCGCGCGCAAGCTGACGCTGGAGCGCATCCGC 66  
QY 61 CTGCGCCCGCGCGCAAGAGTGTACATGATGAAGCACTGTGTGGCCAGCCGCGAG 120  
Db 67 CTGCGCCCGCGCGCAAGAGTGTACATGATGAAGCACTGTGTGGCCAGCCGCGAG 126  
QY 121 CTGAGAAAGTTCGCTTGAACCCCGCGCTGTGAGACCAAGAGGGCTGCAAGCATC 180  
Db 127 CTGAGAAAGTTCGCTTGAACCCCGCGCTGTGAGACCAAGAGGGCTGCAAGCATC 186  
QY 181 ATCCGCAAGCTGCAACCCCGCTGCAAGCCGCGAGAGAGCTGAAGAGCTGTCAAC 240  
Db 187 ATCCGCAAGCTGCAACCCCGCTGCAAGCCGCGAGAGAGCTGAAGAGCTGTCAAC 246  
QY 241 ACCGTGGCCACCTGTACTGTGTCAGAGAGATGAGAGTCCGCAACCAAGAGGCC 300  
Db 247 ACCGTGGCCACCTGTACTGTGTCAGAGAGATGAGAGTCCGCAACCAAGAGGCC 306  
QY 301 CTGACAAAGATGAG 360  
Db 307 CTGACAAAGATGAG 366  
QY 361 GCCGCAAGGGCAAGGTAGGCCAGAACTACCCCATCTGTGAGAACTGACAGGGCCAGATG 420  
Db 367 GCCGCAAGGGCAAGGTAGGCCAGAACTACCCCATCTGTGAGAACTGACAGGGCCAGATG 426  
QY 421 GTGCAACAGGCGCATCAGCCCGCGCACCTGAAAGCTGGTGAAGGTGATGAGAGAAG 480  
Db 427 GTGCAACAGGCGCATCAGCCCGCGCACCTGAAAGCTGGTGAAGGTGATGAGAGAAG 486  
QY 481 GCCTTCAAGCCCGAGGTGATCCCATATTTCAACCGCCCTGAGCGAGGGCGCCAG 540  
Db 487 GCCTTCAAGCCCGAGGTGATCCCATATTTCAACCGCCCTGAGCGAGGGCGCCAG 546  
QY 541 GACTGAACAGATGTTGAACCCGTGGCGCCACAGGCGCCCATGACAGTGTGAAG 600  
Db 547 GACTGAACAGATGTTGAACCCGTGGCGCCACAGGCGCCCATGACAGTGTGAAG 606  
QY 601 GACACCATCAACGAGAGCGCGCGAGTGGACCGCGTGCACCCCGTGCACCGCGGCCCC 660  
Db 607 GACACCATCAACGAGAGCGCGCGAGTGGACCGCGTGCACCCCGTGCACCGCGGCCCC 666  
QY 661 ATGCCCCCGCCAGATGCGGAGAGCCCGCGCGAGCAATCGCCGCGCACCAAGCACCC 720  
Db 667 ATGCCCCCGCCAGATGCGGAGAGCCCGCGCGAGCAATCGCCGCGCACCAAGCACCC 726  
QY 721 CTGCAAGAGAGATGCGCTGTGATGACCAAGCAACCCCGCATCCCGTGGCGCATCTAC 780  
Db 727 CTGCAAGAGAGATGCGCTGTGATGACCAAGCAACCCCGCATCCCGTGGCGCATCTAC 786  
QY 781 AAGCGGTGATCATCTGTGGCTTGAACAAGATCTGCGATGTACAGCCCGGTGAGCATC 840  
Db 787 AAGCGGTGATCATCTGTGGCTTGAACAAGATCTGCGATGTACAGCCCGGTGAGCATC 846  
QY 841 CTGACATCAAGCAGGGGCCCAAGAGCCCTTCCGACTACGTGAGCCGCTTCTTCAAG 900  
Db 847 CTGACATCAAGCAGGGGCCCAAGAGCCCTTCCGACTACGTGAGCCGCTTCTTCAAG 906  
QY 901 ACCCTGCGCGCCGAGCAGAGCAACCCAGAGGTGAAGAACTGTGATGACCCGACACCTGCTG 960  
Db 907 ACCCTGCGCGCCGAGCAGAGCAACCCAGAGGTGAAGAACTGTGATGACCCGACACCTGCTG 966  
QY 961 GTGACAGACGCCAACCCCGACTGCAAGACCATCTGCGCGCTCTGCGCCCGCGCGCAGC 1020

DB 967 GTGCAAGAACCCCAACCCGACTGCAAGACCATTCTGCGGCTCTCGGCCCCGCGCCAGC 1026  
QY 1021 CTGAGAGAGATGATGACCGCTCTGCCAGGGCGTGGCGGCCCCCAAGCCCAAGGCCGCTG 1080  
DB 1027 CTGAGAGAGATGATGACCGCTCTGCCAGGGCGTGGCGGCCCCCAAGGCCGCTG 1086  
QY 1081 CTGGCCGAGGAGTGAAGCCAGGCCAACAACAGGCTGATGATGCAAGAGCACTTCAAG 1140  
DB 1087 CTGGCCGAGGAGTGAAGCCAGGCCAACAACAGGCTGATGATGCAAGAGCACTTCAAG 1146  
QY 1141 GGGCCCCGCGCATGCTCAAGTCTTCAACTGGCCAGAGAGGSCCAATCGCCCGCAAC 1200  
DB 1147 GGGCCCCGCGCATGCTCAAGTCTTCAACTGGCCAGAGAGGSCCAATCGCCCGCAAC 1206  
QY 1201 TGCCGCGCCCCCGCAAGAGGCTGTGGAAGTGCAGGCAAGAGGGCCCAACAGATGAAG 1260  
DB 1207 TGCCGCGCCCCCGCAAGAGGCTGTGGAAGTGCAGGCAAGAGGGCCCAACAGATGAAG 1266  
QY 1261 GACTGCAACCGAGCGCCAGGCCAATTCTGGGCAAGATCTGGCCCAAGAGGGCCGC 1320  
DB 1267 GACTGCAACCGAGCGCCAGGCCAATTCTGGGCAAGATCTGGCCCAAGAGGGCCGC 1326  
QY 1321 CCGCGCACTTCTCTGCAAGAGCGGCCCCGAGCCACCGCCCCCGCGAGAGCTTCCGC 1380  
DB 1327 CCGCGCACTTCTCTGCAAGAGCGGCCCCGAGCCACCGCCCCCGCGAGAGCTTCCGC 1386  
QY 1381 TTCGAGAGAGCAACACCCCGCGCAAGAGCAAGAGCAAGAGACCGAGACCTGAACAGC 1440  
DB 1387 TTCGAGAGAGCAACACCCCGCGCAAGAGCAAGAGCAAGAGACCGAGACCTGAACAGC 1446  
QY 1441 CTGAGAGAGCTGTCTGCGCAAGCAACCCCTGAGCCAGTAA 1479  
DB 1447 CTGAGAGAGCTGTCTGCGCAAGCAACCCCTGAGCCAGTAA 1485

RESULT 6  
US-10-190-305A-14

Sequence 14, Application US/10190305A  
Publication No. US20030198621A1  
GENERAL INFORMATION:  
APPLICANT: ZUR MEGEDR, Jan  
APPLICANT: BARNETT, Susan  
APPLICANT: LIAN, Ying  
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B AND/OR  
TITLE OF INVENTION: TYPE C POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF  
FILE REFERENCE: 2302-18702 / 18702.002  
CURRENT APPLICATION NUMBER: US/10/190,305A  
CURRENT FILING DATE: 2002-07-05  
NUMBER OF SEQ ID NOS: 93  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 14  
LENGTH: 4419  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURES:  
OTHER INFORMATION: Description of Artificial Sequence:  
US-10-190-305A-14

Query Match 99.9%; Score 1477.4; DB 15; Length 4419;  
Best Local Similarity 99.9%; Pred. No. 1.4e-302;  
Matches 1478; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGCGCGCCAGCATCTGCGCGCGCGCAAGCTGAGCCCTGGAGCGCATCCGC 60  
DB 7 ATGGGCGCGCGCCAGCATCTGCGCGCGCGCAAGCTGAGCCCTGGAGCGCATCCGC 66  
QY 61 CTGCGCGCGCGCGCAAGAGTCTACATGATGAAGCACTGTGTGGGCGCAGCCGCGAG 120  
DB 67 CTGCGCGCGCGCGCAAGAGTCTACATGATGAAGCACTGTGTGGGCGCAGCCGCGAG 126  
QY 121 CTGGAAGAGTTGCGCTGAACCCCGGCTGTGAGAGCAAGCGAGGCTGCAAGCAGATC 180

DB 127 CTGGAAGAGTTGCGCTGAACCCCGGCTGTGAGAGCAAGCGAGGCTGCAAGCAGATC 186  
QY 181 ATCCGCAAGTCAACCCCGGCTGCAAGACCGGCAAGAGAGCTGAAGAGCTGTCAAC 240  
DB 187 ATCCGCAAGTCAACCCCGGCTGCAAGACCGGCAAGAGAGAGCTGTCAAC 246  
QY 241 ACCGTGGCCACCTGTACTGCTGTCACGAGAGATGAGGTCCGCGACCAAGAGGCGC 300  
DB 247 ACCGTGGCCACCTGTACTGCTGTCACGAGAGATGAGGTCCGCGACCAAGAGGCGC 306  
QY 301 CTGGAAGAATGCAAGAGAGAGCAACAAGTCCAGAGATCCAGAGGCGAGGCGC 360  
DB 307 CTGGAAGAATGCAAGAGAGAGCAACAAGTCCAGAGATCCAGAGGCGAGGCGC 366  
QY 361 GCGGCAAGGCGCAAGTGAAGCAGAACTACCCCATGTGCAAGACTGCAAGGCGCAGATG 420  
DB 367 GCGGCAAGGCGCAAGTGAAGCAGAACTACCCCATGTGCAAGACTGCAAGGCGCAGATG 426  
QY 421 GTGCAACAGGCGCATCAGCCCCCGCAACCTGAAAGCTGTGAGAGTGAAGAGAGAG 480  
DB 427 GTGCAACAGGCGCATCAGCCCCCGCAACCTGAAAGCTGTGAGAGTGAAGAGAGAG 486  
QY 481 GCTTTCAGCCCCGAGGTGATTCCTCATGTTCACCGCTGAGCGGCGGCGCAACCCCCAG 540  
DB 487 GCTTTCAGCCCCGAGGTGATTCCTCATGTTCACCGCTGAGCGGCGGCGCAACCCCCAG 546  
QY 541 GACTTGAACAGATGTTGAACACCTGTGGCGGCGCAACAGGCGGCGCATGCAAGTGTGAAG 600  
DB 547 GACTTGAACAGATGTTGAACACCTGTGGCGGCGCAACAGGCGGCGCATGCAAGTGTGAAG 606  
QY 601 GACACCATCAAGAGAGAGGCGCGCGAGTGGACCGCGTGCAACCCCGTGCAACGCGGCGCC 660  
DB 607 GACACCATCAAGAGAGAGGCGCGCGAGTGGACCGCGTGCAACCCCGTGCAACGCGGCGCC 666  
QY 661 ATGCGCCCCCGGCGAGATGCGCGAGACCCCGCGGCGAGCATTCGCGCGCACCAACCAACC 720  
DB 667 ATGCGCCCCCGGCGAGATGCGCGAGACCCCGCGGCGAGCATTCGCGCGCACCAACCAACC 726  
QY 721 CTGCAAGAGCAGATGCGCTGTGATGACCAAGCAACCCCATTCCTGCGGCGAGATCTAC 780  
DB 727 CTGCAAGAGCAGATGCGCTGTGATGACCAAGCAACCCCATTCCTGCGGCGAGATCTAC 786  
QY 781 AAGCGGTGATCATCTGGGCTGAAACAAGATGTGCGGATGTACAGCCCGTGAGCATC 840  
DB 787 AAGCGGTGATCATCTGGGCTGAAACAAGATGTGCGGATGTACAGCCCGTGAGCATC 846  
QY 841 CTGGAATCAAGCAGGCGCGCGCAAGAGGCGCTTCGCGACTACGTGAGCGCTTCTTCAAG 900  
DB 847 CTGGAATCAAGCAGGCGCGCGCAAGAGGCGCTTCGCGACTACGTGAGCGCTTCTTCAAG 906  
QY 901 ACCCTGCGCGCGAGCAGAGCAACCCAGAGGTGAAGAACTGATGACCGACACCTGTCTG 960  
DB 907 ACCCTGCGCGCGAGCAGAGCAACCCAGAGGTGAAGAACTGATGACCGACACCTGTCTG 966  
QY 961 GTGCAAGACGCAACCCCGACTGCAAGAACCATCTGCGCGCTCTGCGCGCGCGCGCGAGC 1020  
DB 967 GTGCAAGACGCAACCCCGACTGCAAGAACCATCTGCGCGCTCTGCGCGCGCGCGCGAGC 1026  
QY 1021 CTGAGAGAGATGATGACCGCTGCAAGGCGCGTGGCGGCGCGCGCGCGCGCGCGT 1080  
DB 1027 CTGAGAGAGATGATGACCGCTGCAAGGCGCGTGGCGGCGCGCGCGCGCGCGCGT 1086  
QY 1081 CTGGCCGAGGCGATGAGCCAGGCCAACAACAGGCTGATGATGCAAGAGCAACTTCAAG 1140  
DB 1087 CTGGCCGAGGCGATGAGCCAGGCCAACAACAGGCTGATGATGCAAGAGCAACTTCAAG 1146  
QY 1141 GGGCCCCGCGCATGCTCAAGTCTTCAACTGGCCAGAGAGGSCCAATCGCCCGCAAC 1200  
DB 1147 GGGCCCCGCGCATGCTCAAGTCTTCAACTGGCCAGAGAGGSCCAATCGCCCGCAAC 1206  
QY 1201 TGCCGCGCCCCCGCAAGAGGCTGTGGAAGTGCAGGCAAGAGGGCCCAACAGATGAAG 1260  
DB 1207 TGCCGCGCCCCCGCAAGAGGCTGTGGAAGTGCAGGCAAGAGGGCCCAACAGATGAAG 1266



OY	1261	GACTGCACCCGAGCGGCCAGGCCAACTTCCTGGGCAAGATCTGGGCCAGGCCACAAGGGCCGC	1320
Db	1267	GACTGCACCCGAGCGGCCAGGCCAACTTCCTGGGCAAGATCTGGGCCAGGCCACAAGGGCCGC	1326
OY	1321	CCCCGCCAATTCTCTGCAGAGCGCGCCCGGAGCCACCGCCCCCCCCTGGAGAAGCTTCCGC	1380
Db	1327	CCCCGCCAATTCTCTGCAGAGCGCGCCCGGAGCCACCGCCCCCCCCTGGAGAAGCTTCCGC	1386
OY	1381	TTGAGGAGAACACCCCCCGGCCAGAAAGCAGAGAGACAAGACC CGAGACCTTGACCAGC	1440
Db	1387	TTGAGGAGAACACCCCCCGGCCAGAAAGCAGAGAGACAAGACC CGAGACCTTGACCAGC	1446
OY	1441	CTGAAGAGCCTGTTCGGCAACGACCCCTGAGCCAGTAA	1479
Db	1447	CTGAAGAGCCTGTTCGGCAACGACCCCTGAGCCAGAAA	1485

## RESULT 7

```

US-10-190-435-35
; Sequence 35, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Retrellita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 35 `
; LENGTH: 4483
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: gp160mod.TV1.dv1V2-gagmod.BW965
US-10-190-435-35

```

Query Match	99.98;	Score 1477.4;	DB 15;	Length 4483;
-------------	--------	---------------	--------	--------------

Query match:	99.99%	Pred. No. 1.4e-302;
Best Local Similarity	99.94%	

Matches 1478; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	ATGGGCGCCCCGCGCCAGCATCTCTGCGCGGCGGCAAGCTGGAGCGCTTGGAGCGCATCCGC	60
Db	3005	ATGGGCGCCCCGCGCCAGCATCTCTGCGCGGCGGCAAGCTGGAGCGCTTGGAGCGCATCCGC	3064
Qy	61	CTGCGCCCCGCGCGGCAAGAGTGTCTACATGATGAAGCACTTGTTGTGGCCAGCCGCGAG	120
Db	3065	CTGCGCCCCGCGCGGCAAGAGTGTCTACATGATGAAGCACTTGTTGTGGCCAGCCGCGAG	3124
Qy	121	CTGAGAAAGTTTCGCCCTGTAAACCCCGGCTGTGCGAGACCAGCGAAGGGCTGCAACCAGATC	180
Db	3125	CTGAGAAAGTTTCGCCCTGTAAACCCCGGCTGTGCGAGACCAGCGAAGGGCTGCAACCAGATC	3184
Qy	181	ATCCGCCAGCTGCACCCCGCCCTGCAGACCCGCGCAGAGAGCTGAAGAGCCTGTTCAAC	240
Db	3185	ATCCGCCAGCTGCACCCCGCCCTGCAGACCCGCGCAGAGAGCTGAAGAGCCTGTTCAAC	3244
Qy	241	ACCGTGGCCACCCTGTACTGCGGTGCACGAGAAAGATCGAGGTCGCGCACACCAAGGAGGCC	300
Db	3245	ACCGTGGCCACCCTGTACTGCGGTGCACGAGAAAGATCGAGGTCGCGCACACCAAGGAGGCC	3304
Qy	301	CTGACAAAGATCGAGGAGGAGCAAGAACAAAGTGCAGCAGAGATCCAGCAGGCTCGAGGCC	360
Db	3305	CTGACAAAGATCGAGGAGGAGCAAGAACAAAGTGCAGCAGAGATCCAGCAGGCTCGAGGCC	3364
Qy	361	GCCGACAAAGGCGAAGGTGAGCCAGAACTACCCCATCTGTGCAGAACCTGCAGGGCCAGATG	420

Db	3365	GCCGACAAAGGGCAAGGTGAGCCAGAACTACCCCATCTGTGCAAACTGTGCAAGGCGCAAGT	3424
QY	421	GTGCACCAAGGCCATCAGCCCCCGCACCCCTGAACGCTTGAGTGTATCGAGGAGAG	480
Db	3425	GTGCACCAAGGCCATCAGCCCCCGCACCCCTGAACGCTTGAGTGTATCGAGGAGAG	3484
QY	481	GCCTTCAGCCCCCGAGGTATCCCATATGTTCAACGCGCTGAGCGAGGGCGCCACCCCGAG	540
Db	3485	GCCTTCAGCCCCCGAGGTATCCCATATGTTCAACGCGCTGAGCGAGGGCGCCACCCCGAG	3544
QY	541	GACCTGAACACGATGTTGAACAACGTGGGCGGCAACAGGCGGCCATGCAGATGCTGAAG	600
Db	3545	GACCTGAACACGATGTTGAACAACGTGGGCGGCAACAGGCGGCCATGCAGATGCTGAAG	3604
QY	601	GACACCATCAACGAGAGGCGCGCGAGTGGGACCGCGTGCAACCCGTGCAACGCGGCCCC	660
Db	3605	GACACCATCAACGAGAGGCGCGCGAGTGGGACCGCGTGCAACCCGTGCAACGCGGCCCC	3664
QY	661	ATCGCCCCCGGCAGATGCGCGAGCGCCCGCGGCAAGCATGCGCGGCAACCAACAGCAAC	720
Db	3665	ATCGCCCCCGGCAGATGCGCGAGCGCCCGCGGCAAGCATGCGCGGCAACCAACAGCAAC	3724
QY	721	CTGCAGAGCAGATCGCCTGGATGACCAACAACCCCATCCCGCTGGGCGCACTCTAC	780
Db	3725	CTGCAGAGCAGATCGCCTGGATGACCAACAACCCCATCCCGCTGGGCGCACTCTAC	3784
QY	781	AAGCGGTGATCATCTGCGGCTGAACAAGATCTGCGGATGAACAACCCCGTGAGCATC	840
Db	3785	AAGCGGTGATCATCTGCGGCTGAACAAGATCTGCGGATGAACAACCCCGTGAGCATC	3844
QY	841	CTGCATCATCAAGCAGGCGCCCAAGGAGCCCTTCGCGCACTACGTTGAACCGCTTCAAG	900
Db	3845	CTGCATCATCAAGCAGGCGCCCAAGGAGCCCTTCGCGCACTACGTTGAACCGCTTCAAG	3904
QY	901	ACCTTCGCGCGCGAGCAAGCAACCAAGAGGTGAAGAACTGGAAGAAGCAACCGCTCTG	960
Db	3905	ACCTTCGCGCGCGAGCAAGCAACCAAGAGGTGAAGAACTGGAAGAAGCAACCGCTCTG	3964
QY	961	GTGCAGAACGCCAACCCCGACTGCAAGACATCTGCGCGCTCTCGGCCCGGCGCAGC	1020
Db	3965	GTGCAGAACGCCAACCCCGACTGCAAGACATCTGCGCGCTCTCGGCCCGGCGCAGC	4024
QY	1021	CTGAGGAGATGATGACCGCTGCAAGGCGTGGGCGGCGCCCAAGCGCGGTG	1080
Db	4025	CTGAGGAGATGATGACCGCTGCAAGGCGTGGGCGGCGCCCAAGCGCGGTG	4084
QY	1081	CTGCGCGAGGCGGATGAGCCAGGCCAACACAGCGTGATGATGAGAAGCACTTCAAG	1140
Db	4085	CTGCGCGAGGCGGATGAGCCAGGCCAACACAGCGTGATGATGAGAAGCACTTCAAG	4144
QY	1141	GGCCCCCGCGCATCTGTTCAAGTGTCTTCAACTGCGGCAAGGAGGCGCATCTGCGCAAC	1200
Db	4145	GGCCCCCGCGCATCTGTTCAAGTGTCTTCAACTGCGGCAAGGAGGCGCATCTGCGCAAC	4204
QY	1201	TGCGCGCGCCCCCGCAAGAGGGCTGTGGAAGTGCGGCAAGGAGGCGCACAGATGAAG	1260
Db	4205	TGCGCGCGCCCCCGCAAGAGGGCTGTGGAAGTGCGGCAAGGAGGCGCACAGATGAAG	4264
QY	1261	GACTGCAACCGAGCGCGCAAGCTTCTCTGGGCAAGATCTGCGCAAGGAGGCGCGC	1320
Db	4265	GACTGCAACCGAGCGCGCAAGCTTCTCTGGGCAAGATCTGCGCAAGGAGGCGCGC	4324
QY	1321	CCCGGCAACTTCTGAGAGCGCGCCGAGCCCAACCGCGCCCGCGGAGAGCTTCCGC	1380
Db	4325	CCCGGCAACTTCTGAGAGCGCGCCGAGCCCAACCGCGCCCGCGGAGAGCTTCCGC	4384
QY	1381	TTTGAAGAGACCAACCCCGGCGAAGGCAAGAGCAAGACCGGAGACCTGTACAGC	1440
Db	4385	TTTGAAGAGACCAACCCCGGCGAAGGCAAGAGCAAGACCGGAGACCTGTACAGC	4444
QY	1441	CTGAAGAGCCTGTTGCGCAACGAGCCCTGTAGCCAGTAA	1479



Db 4445 CTGAAGAGCTGTTCCGCAACGACCCCTGAGCCATTA 4483

## RESULT 8

US-10-190-435-34

; Sequence 34, Application US/10190435

; Publication No. US20030143248A1

; GENERAL INFORMATION:

; APPLICANT: ZUR MEGEDE, Jan

; APPLICANT: BARNETT, Susan W.

; APPLICANT: LIAN, Ying

; APPLICANT: ENGELBRECHT, Susan

; APPLICANT: VAN RENSBURG, Estrelita J.

; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C

; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF

; FILE REFERENCE: P18133.003 / 2302-18133

; CURRENT APPLICATION NUMBER: US/10/190,435

; NUMBER OF SEQ ID NOS: 319

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 34

; LENGTH: 4606

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURES:

; OTHER INFORMATION: Description of Artificial Sequence:

; OTHER INFORMATION: gp160mod.TV1.dvl-gagmod.BW965

; US-10-190-435-34

Query Match 99.9%; Score 1477.4; DB 15; Length 4606;

Best Local Similarity 99.9%; Pred. No. 1.4e-302;

Matches 1478; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGCGCCCGCCAGCATCTCTCGCGCGCGCGCAAGCTGACCGCTGGAGCGCATCCG 60  
DB 3128 ATGGCGCGCCCGCCAGCATCTCTCGCGCGCGCGCAAGCTGACCGCTGGAGCGCATCCG 3187  
QY 61 CTGGCGCGCCCGCCAGCAAGAGTGTACATGATGAGCACTGTGTGGGCGCAGCGCGAG 120  
DB 3188 CTGGCGCGCCCGCCAGCAAGAGTGTACATGATGAGCACTGTGTGGGCGCAGCGCGAG 3247  
QY 121 CTGGAAGAGTTCCTCTGAAACCCCGCTCTCTGAGACCAAGCGGCTGCAAGAGATC 180  
DB 3248 CTGGAAGAGTTCCTCTGAAACCCCGCTCTCTGAGACCAAGCGGCTGCAAGAGATC 3307  
QY 181 ATCCGCCAGCTGCAACCCCGCTCTGAGACCGGCGGAGAGCTGAAGAGCTTTCAAC 240  
DB 3308 ATCCGCCAGCTGCAACCCCGCTCTGAGACCGGCGGAGAGCTGAAGAGCTTTCAAC 3367  
QY 241 ACCGTGGCCACCTCTGTACTGCTGTGCAAGAGAGTGAAGTCCGCGACACCAAGAGGCC 300  
DB 3368 ACCGTGGCCACCTCTGTACTGCTGTGCAAGAGAGTGAAGTCCGCGACACCAAGAGGCC 3427  
QY 301 CTGGAAGAGTTCGAG 360  
DB 3428 CTGGAAGAGTTCGAG 3487  
QY 361 GCCGACAAAGGCAAGTGAAGCAAACTACCCCATCTGTGAGAACTTCAGAGGCCAGATG 420  
DB 3488 GCCGACAAAGGCAAGTGAAGCAAACTACCCCATCTGTGAGAACTTCAGAGGCCAGATG 3547  
QY 421 GTGCAACCAAGGCAATCAAGCCCGCCGACCTGAAACGCTGGGTGAAGGTGATGAGAGAA 480  
DB 3548 GTGCAACCAAGGCAATCAAGCCCGCCGACCTGAAACGCTGGGTGAAGGTGATGAGAGAA 3607  
QY 481 GCCTTCAAGCCCGCAAGTATCCCATGTTCAACCGCTCTGAGCGAGAGGCCACCCCCAG 540  
DB 3608 GCCTTCAAGCCCGCAAGTATCCCATGTTCAACCGCTCTGAGCGAGAGGCCACCCCCAG 3667  
QY 541 GACCTGAACAGATGTTGAACACCGTGGGCGGCGACCAAGGCCCATGAGATGCTGAAG 600  
DB 3668 GACCTGAACAGATGTTGAACACCGTGGGCGGCGACCAAGGCCCATGAGATGCTGAAG 3727

QY 601 GACACCATCAAGAGAGAGCGCGGAGTGGAGCCGCTGCAACCCCGTGACAGCGCGCCCC 660  
DB 3728 GACACCATCAAGAGAGAGCGCGGAGTGGAGCCGCTGCAACCCCGTGACAGCGCGCCCC 3787  
QY 661 ATCGCCCCCGGCAAGTATGCGGAGAGCGCGCGGAGGAGATGCGCGGACCAAGAGACC 720  
DB 3788 ATCGCCCCCGGCAAGTATGCGGAGAGCGCGCGGAGGAGATGCGCGGACCAAGAGACC 3847  
QY 721 CTGCAAGAGAGATGCGCTGTGATGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780  
DB 3848 CTGCAAGAGAGATGCGCTGTGATGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3907  
QY 781 AAGCGGTGATCATCTGCGGCTGAACAAGATGCGGATGATGACAGCCCGTGAGCATC 840  
DB 3908 AAGCGGTGATCATCTGCGGCTGAACAAGATGCGGATGATGACAGCCCGTGAGCATC 3967  
QY 841 CTGCAATCAAG 900  
DB 3968 CTGCAATCAAG 4027  
QY 901 ACCCTGCGCGCGGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960  
DB 4028 ACCCTGCGCGCGGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4087  
QY 961 CTGCAAG 1020  
DB 4088 CTGCAAG 4147  
QY 1021 CTGGAAGAGATGATGACCGCTTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080  
DB 4148 CTGGAAGAGATGATGACCGCTTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4207  
QY 1081 CTGCGCGAG 1140  
DB 4208 CTGCGCGAG 4267  
QY 1141 GCGCCCCCGGCGCATGCTCAAGTCTTCAACTGCGGCAAGAGAGAGAGAGAGAGAGAG 1200  
DB 4268 GCGCCCCCGGCGCATGCTCAAGTCTTCAACTGCGGCAAGAGAGAGAGAGAGAGAGAG 4327  
QY 1201 TGCGCGCGCGCGCGCAAG 1260  
DB 4328 TGCGCGCGCGCGCGCAAG 4387  
QY 1261 GACTGCAACGAGCGGCAAG 1320  
DB 4388 GACTGCAACGAGCGGCAAG 4447  
QY 1321 CCGGCAACTTCTCTGCAAGAGCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380  
DB 4448 CCGGCAACTTCTCTGCAAGAGCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4507  
QY 1381 TTGAG 1440  
DB 4508 TTGAG 4567  
QY 1441 CTGAAGAGCTGTTGCGCAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1479  
DB 4568 CTGAAGAGCTGTTGCGCAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4606

## RESULT 9

US-10-190-435-36

; Sequence 36, Application US/10190435

; Publication No. US20030143248A1

; GENERAL INFORMATION:

; APPLICANT: ZUR MEGEDE, Jan

; APPLICANT: BARNETT, Susan W.

; APPLICANT: LIAN, Ying

; APPLICANT: ENGELBRECHT, Susan

; APPLICANT: VAN RENSBURG, Estrelita J.

; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C

; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF

FILE REFERENCE: PP18133.003 / 2302-18133  
CURRENT APPLICATION NUMBER: US/10/190,435  
CURRENT FILING DATE: 2002-12-30  
NUMBER OF SEQ ID NOS: 319  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 36  
LENGTH: 4615  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence:  
US-10-190-435-36

Query Match 99.9%; Score 1477.4; DB 15; Length 4615;  
Best Local Similarity 99.9%; Pred. No. 1.4e-302;  
Matches 1478; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGCCCGCCAGCATCTGCGCGCGCGCAAGCTGACCCCTGGAGCGCATCCGC 60  
DB 3137 ATGGGCGCCCGCCAGCATCTGCGCGCGCGCAAGCTGACCCCTGGAGCGCATCCGC 3196  
QY 61 CTGCGCGCCCGCGCAAGAGTCTACATGATGAAGCACCTGTGTGGCCAGCCGCGAG 120  
DB 3197 CTGCGCGCCCGCGCAAGAGTCTACATGATGAAGCACCTGTGTGGCCAGCCGCGAG 3256  
QY 121 CTGAGAGAGTTCGCTGAAACCCCGCGCTGCTGAGAGCAGCGAGGCTGCAAGCAGATC 180  
DB 3257 CTGAGAGAGTTCGCTGAAACCCCGCGCTGCTGAGAGCAGCGAGGCTGCAAGCAGATC 3316  
QY 181 ATCCGCGAGCTGACCCCGCGCTGAGACCGCGCAGAGAGCTGAAGAGCTGTTCAAC 240  
DB 3317 ATCCGCGAGCTGACCCCGCGCTGAGACCGCGCAGAGAGCTGTTCAAC 3376  
QY 241 ACCGTGGCACTCTGTACTGCGTGCACGAGAGATCGAGGTCCGCGACACCAAGAGGCC 300  
DB 3377 ACCGTGGCACTCTGTACTGCGTGCACGAGAGATCGAGGTCCGCGACACCAAGAGGCC 3436  
QY 301 CTGGAACAAGTCGAGGAGGAGCAACAAGTCCAGCAAGATCCAGAGCGCGAGGCC 360  
DB 3437 CTGGAACAAGTCGAGGAGGAGCAACAAGTCCAGCAAGATCCAGAGCGCGAGGCC 3496  
QY 361 GCCGACAAGGCAAGGTGAGGCAAGTACCCCATCTGCAAGAACTGCAAGGCGCAGATG 420  
DB 3497 GCCGACAAGGCAAGGTGAGGCAAGTACCCCATCTGCAAGAACTGCAAGGCGCAGATG 3556  
QY 421 GTGACACAGGCGCATCAGCCCGCGCACCTGAACGCTGGTGAAGGTATCGAGGAAG 480  
DB 3557 GTGACACAGGCGCATCAGCCCGCGCACCTGAACGCTGGTGAAGGTATCGAGGAAG 3616  
QY 481 GCCTTCAAGCCCGAGGTATCCCATGTTTCAACCGCCCTGAGCGAGGCGCGCACCCCCAG 540  
DB 3617 GCCTTCAAGCCCGAGGTATCCCATGTTTCAACCGCCCTGAGCGAGGCGCGCACCCCCAG 3676  
QY 541 GACCTGAACAGATGTTGAACAACGTGGCGCGCACCAAGGCGCGCATGCAATGCTGAAG 600  
DB 3677 GACCTGAACAGATGTTGAACAACGTGGCGCGCACCAAGGCGCGCATGCAATGCTGAAG 3736  
QY 601 GACACCATCAACGAGGAGGCGCGCGAGTGGGACCGGTGCAACCCCGTGCACGCGCGGCC 660  
DB 3737 GACACCATCAACGAGGAGGCGCGCGAGTGGGACCGGTGCAACCCCGTGCACGCGCGGCC 3796  
QY 661 ATGCGCGCGCGCGCAGATGCGCGAGAGCCCGCGCGCAGGACATGCGCGCGCACCAAGC 720  
DB 3797 ATGCGCGCGCGCGCAGATGCGCGAGAGCCCGCGCGCAGGACATGCGCGCGCACCAAGC 3856  
QY 721 CTGCAAGAGCAGATGCGCTGATGACCAAGAACCCCGCATCCCGTGGCGCATCTAC 780  
DB 3857 CTGCAAGAGCAGATGCGCTGATGACCAAGAACCCCGCATCCCGTGGCGCATCTAC 3916  
QY 781 AAGCGGTGATCATCTTGGGCTTGAACAAGATCTGCGGATGTAACAGCCCGGTGAGCATC 840  
DB 3917 AAGCGGTGATCATCTTGGGCTTGAACAAGATCTGCGGATGTAACAGCCCGGTGAGCATC 3976

QY 841 CTGACATCAAGCAGGCGCCCAAGAGAGCCCTTCGCGACTACGTGACCGCTTCTCAAG 900  
DB 3977 CTGACATCAAGCAGGCGCCCAAGAGAGCCCTTCGCGACTACGTGACCGCTTCTCAAG 4036  
QY 901 ACCCTGCGCGCGCAGAGCAGCAGCAGAGGTGAAGACTGTGATACCGACACCTGCTG 960  
DB 4037 ACCCTGCGCGCGCAGAGCAGCAGCAGAGGTGAAGACTGTGATACCGACACCTGCTG 4096  
QY 961 GTGCAAGACGCCAACCCCGACTGCAAGAACCATCTCGCGGCTCTGCGCCCGCGCGAGC 1020  
DB 4097 GTGCAAGACGCCAACCCCGACTGCAAGAACCATCTCGCGGCTCTGCGCCCGCGCGAGC 4156  
QY 1021 CTGAGAGAGATGATGACCGCTGCAAGGCGCGCGCGCGCGCGCGCGCGCGCGCG 1080  
DB 4157 CTGAGAGAGATGATGACCGCTGCAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4216  
QY 1081 CTGCGCGAGGCGCATGAGCCAGCGCAACACCAAGCTGATGATGCAAGAGCAACTCAAG 1140  
DB 4217 CTGCGCGAGGCGCATGAGCCAGCGCAACACCAAGCTGATGATGCAAGAGCAACTCAAG 4276  
QY 1141 GCGCCCGCGCGCATCTGCAAGTCTTCAACTGCGCGCAAGAGGCGCAACATCGCCCGCAAC 1200  
DB 4277 GCGCCCGCGCGCATCTGCAAGTCTTCAACTGCGCGCAAGAGGCGCAACATCGCCCGCAAC 4336  
QY 1201 TGCCTGCGCGCGCGCGCAAGAGGCGCTGGAAGTGCAGAGGCGCGCAAGAGGCGCGCG 1260  
DB 4337 TGCCTGCGCGCGCGCGCAAGAGGCGCTGGAAGTGCAGAGGCGCGCAAGAGGCGCGCG 4396  
QY 1261 GACTGCACCGAGCGCGCAGCGCAACTTCTGCGCGCAAGATCTGCGCGCAAGAGCGCGCGCG 1320  
DB 4397 GACTGCACCGAGCGCGCAGCGCAACTTCTGCGCGCAAGATCTGCGCGCAAGAGCGCGCGCG 4456  
QY 1321 CCGCGCAACTTCTGTCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380  
DB 4457 CCGCGCAACTTCTGTCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4516  
QY 1381 TTCGAGGAGACCAACCCCGCGCGCAAGAGCAGAGAGCAAGAGCGCGCGCGCGCGCGCG 1440  
DB 4517 TTCGAGGAGACCAACCCCGCGCGCAAGAGCAGAGAGCAAGAGCGCGCGCGCGCGCGCG 4576  
QY 1441 CTGAAGAGCTGTTTGGCAAGCAGACCCCTGAGCGCACTAA 1479  
DB 4577 CTGAAGAGCTGTTTGGCAAGCAGACCCCTGAGCGCACTAA 4615

## RESULT 10

US-10-190-435-38

; Sequence 38, Application US/10190435

; Publication No. US2003014324BA1

; GENERAL INFORMATION:

; APPLICANT: ZUR MEGEDR, Jan

; APPLICANT: BARNETT, Susan W.

; APPLICANT: LIAN, Yang

; APPLICANT: ENGELBRECHT, Susan

; APPLICANT: VAN RENSBURG, Estrelita J.

; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C

; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF

; FILE REFERENCE: PP18133.003 / 2302-18133

; CURRENT APPLICATION NUMBER: US/10/190,435

; CURRENT FILING DATE: 2002-12-30

; NUMBER OF SEQ ID NOS: 319

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 38

; LENGTH: 4702

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence:

; US-10-190-435-38

Query Match 99.9%; Score 1477.4; DB 15; Length 4702;

Best Local Similarity 99.9%; Pred. No. 1,4e-302;  
Matches 1478; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	ATGGGCGCCGCGCAGCATCTCTGCGCGCGCAAGCTTGGACGCTTGGAGCGCATCCG	60
DB	3224	ATGGGCGCCGCGCAGCATCTCTGCGCGCGCAAGCTTGGACGCTTGGAGCGCATCCG	3283
QY	61	CTGCGCCCCGCGCGCAAGAGTGTCTACATGATGAAGCACTTGTGTGTGGCCAGCCGCGAG	120
DB	3284	CTGCGCCCCGCGCGCAAGAGTGTCTACATGATGAAGCACTTGTGTGTGGCCAGCCGCGAG	3343
QY	121	CTGGAAGAGTTTCGCTCTGAACCCCGGCTGTCTGTGAAGACCAAGGAGGGCTGTGAAGCATC	180
DB	3344	CTGGAAGAGTTTCGCTCTGAACCCCGGCTGTCTGTGAAGACCAAGGAGGGCTGTGAAGCATC	3403
QY	181	ATTCGCGCAGCTGCACCCCGGCTGTGCAGACCGGCAAGGAGAGCTGAAGAGCTGTTCAC	240
DB	3404	ATTCGCGCAGCTGCACCCCGGCTGTGCAGACCGGCAAGGAGAGCTGAAGAGCTGTTCAC	3463
QY	241	ACCGTGGCCACCCCTGTACTGTCTGTGCAGAGAGATCGAGTTCGCGCAACCAAGAGAGGCC	300
DB	3464	ACCGTGGCCACCCCTGTACTGTCTGTGCAGAGAGATCGAGTTCGCGCAACCAAGAGAGGCC	3523
QY	301	CTGGAACAAGATCGAGAGAGAGAGACAACAAGTGCAGCAGAGAGATCCAGCAGGCCAGAGCC	360
DB	3524	CTGGAACAAGATCGAGAGAGAGAGACAACAAGTGCAGCAGAGAGATCCAGCAGGCCAGAGCC	3583
QY	361	GCCGACAAAGGGCAAGGTGAGCCAGAACTACCCCATCTGTGAGAACTGTACAGGGCCAGATG	420
DB	3584	GCCGACAAAGGGCAAGGTGAGCCAGAACTACCCCATCTGTGAGAACTGTACAGGGCCAGATG	3643
QY	421	GTCACACAGGCCATCAGCCCCCGCAACCTGAACGCTGTGGTGAAGGTGATCGAGAGAAAG	480
DB	3644	GTCACACAGGCCATCAGCCCCCGCAACCTGAACGCTGTGGTGAAGGTGATCGAGAGAAAG	3703
QY	481	GCTTTCAGCCCCGAGGTGATCCCATGTTTCAACGCGCTTGAAGAGGGCGGCCACCCCCAG	540
DB	3704	GCTTTCAGCCCCGAGGTGATCCCATGTTTCAACGCGCTTGAAGAGGGCGGCCACCCCCAG	3763
QY	541	GACCTGAAACAAGATGTTGAAACAACGTGGGCGGCGCACAGGCCGCTATGCAGATGCTGAAG	600
DB	3764	GACCTGAAACAAGATGTTGAAACAACGTGGGCGGCGCACAGGCCGCTATGCAGATGCTGAAG	3823
QY	601	GACACCATCAACGAGAGAGGCGCGCAGTGGGACCGCGTGCACCCGCTGCACGCGCGCCCC	660
DB	3824	GACACCATCAACGAGAGAGGCGCGCAGTGGGACCGCGTGCACCCGCTGCACGCGCGCCCC	3883
QY	661	ATCGCCCCCGCCAGATGCGCGCAGCCCCCGCGCAGCGACATCGCCGCAACCAACAGCACCC	720
DB	3884	ATCGCCCCCGCCAGATGCGCGCAGCCCCCGCGCAGCGACATCGCCGCAACCAACAGCACCC	3943
QY	721	CTGCAGAGCAGATCGCCTGTGATGACCAAGCAACCCCCCATCCCGTGGGCGCATCTTAC	780
DB	3944	CTGCAGAGCAGATCGCCTGTGATGACCAAGCAACCCCCCATCCCGTGGGCGCATCTTAC	4003
QY	781	AAGCGGTGATCATCTCTGGGCTTGAAACAAGATCGTCCGAGTACAGTGAACCGCTTCTCAAG	840
DB	4004	AAGCGGTGATCATCTCTGGGCTTGAAACAAGATCGTCCGAGTACAGTGAACCGCTTCTCAAG	4063
QY	841	CTGCACATCAAGAGAGGGCCCCCAAGAGCCCTTCCGCGACTACGTGGAACCGCTTCTCAAG	900
DB	4064	CTGCACATCAAGAGAGGGCCCCCAAGAGCCCTTCCGCGACTACGTGGAACCGCTTCTCAAG	4123
QY	901	ACCGTGCAGCGCCGAGCAGAGCAACCAAGAGGTGAAGAACTGATGACCGACACCCCTGCTG	960
DB	4124	ACCGTGCAGCGCCGAGCAGAGCAACCAAGAGGTGAAGAACTGATGACCGACACCCCTGCTG	4183
QY	961	GTCGACGAACGCCAACCCCGCACTGCAAGACCATCTCGCGCGCTCTCGGCCCCCGCGCCAGC	1020
DB	4184	GTCGACGAACGCCAACCCCGCACTGCAAGACCATCTCGCGCGCTCTCGGCCCCCGCGCCAGC	4243
QY	1021	CTGGAAGAGATGATGACCGCTGTGCAGGGCGTGGGCGGCCCAAGCCCAAGGCCCGCGTG	1080

Db	4244	CTGAGAGAGATGATGATGACCGCCTTGCCAGGCGCTG3CGGCGCCCAAGCCACAAGGCCCGCGTG	4303
Qy	1081	CTGGCCGAGGCGATGAGGCCAGGCCCAACACACAGCGTGATGATGTCAGAGAGCAACTTCAAG	1140
Db	4304	CTGGCCGAGGCGATGAGCCAGGCCCAACACACAGCGTGATGATGTCAGAGAGCAACTTCAAG	4363
Qy	1141	GCCCCCGGCGCATGTCATAGTGCTTCAACTGCGGCAGAGAGGGCCACATTCGCCCGCAAC	1200
Db	4364	GCCCCCGGCGCATGTCATAGTGCTTCAACTGCGGCAGAGAGGGCCACATTCGCCCGCAAC	4423
Qy	1201	TGCGCGCCCCCGCAAGAGGGCTGCTGGAAGTGCGGCGAAGAGGGCCACAGATGAAG	1260
Db	4424	TGCGCGCCCCCGCAAGAGGGCTGCTGGAAGTGCGGCGAAGAGGGCCACAGATGAAG	4483
Qy	1261	GACTGCACCGAGCGCCAGGCCCACTTCTGCGGCAGAGATCTGGCCCAAGCCACAAGGGCGCG	1320
Db	4484	GACTGCACCGAGCGCCAGGCCCACTTCTGCGGCAGAGATCTGGCCCAAGCCACAAGGGCGCG	4543
Qy	1321	CCCGGCMACTTCTGACAGAGCGCGCCCGAGGCCACCGCGCCCCCGCGCGGAGACTTCCGC	1380
Db	4544	CCCGGCMACTTCTGACAGAGCGCGCCCGAGGCCACCGCGCCCCCGCGCGGAGACTTCCGC	4603
Qy	1381	TTTCGAGGAGACCACCCCCGGCCAGAAAGCAGAGAGCAAGGACCGCGAGACCTTGACCAAGC	1440
Db	4604	TTTCGAGGAGACCACCCCCGGCCAGAAAGCAGAGAGCAAGGACCGCGAGACCTTGACCAAGC	4663
Qy	1441	CTGAAAGAGCTGTTCGGCAACGACCCCTTGAGCCAGTAA	1479
Db	4664	CTGAAAGAGCTGTTCGGCAACGACCCCTTGAGCCAGTAA 4702	

```

RESULT 11
US-10-190-435-17
; Sequence 17, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE 1
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 4716
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: GagProteinArtmuttAtRevnef_C
US-10-190-435-17

```

Query Match	99.9%	Score 1477.4	DB 15	Length 4716
Best Local Similarity	99.9%	Pred. No. 1.4e-302		
Matches 1478	Conservative	0	Mismatches 1	Indels 0
			Gaps	0
QY	1 ATGGCGCGCCGCGCCAGCATCTCTGCGCGCGCGCAAGCTGGAACCCCTGGAGCCATCCGC	60		
Db	7 ATGGCGCGCCGCGCCAGCATCTCTGCGCGCGCGCAAGCTGGAACCCCTGGAGCCATCCGC	66		
QY	61 CTGGCCCCCGCGCGCAAGAGTCTACATGATGAAGCACCTGTGTGGGCCAGCCGCGAG	120		
Db	67 CTGGCCCCCGCGCGCAAGAGTCTACATGATGAAGCACCTGTGTGGGCCAGCCGCGAG	126		
QY	121 CTGAGAGAGTTGCGCCCTGAAACCCCGGCTGTGAGAGACCAAGGAGGGCTGCAAGCATC	180		
Db	127 CTGAGAGAGTTGCGCCCTGAAACCCCGGCTGTGAGAGACCAAGGAGGGCTGCAAGCATC	186		





```
427 GTGACACGAGCCATCAGCCCCCGCACCTTGAACGCTGTGTAAGTGTATGAGAGAAAG 486
481 GCCTTACAGCCCGAGGTGATCCCATGTTCACCGCCCTGAGGAGGGCGCCACCCCGCAG 540
487 GCCTTACAGCCCGAGGTGATCCCATGTTCACCGCCCTGAGGAGGGCGCCACCCCGCAG 546
541 GACCTGAACAAGATGTTGAACAACCGTGGCGGCCACCAAGCCCGCATGACATGCTGAAG 600
547 GACCTGAACAAGATGTTGAACAACCGTGGCGGCCACCAAGCCCGCATGACATGCTGAAG 606
601 GACACCATCAACGAGAGGCGCCGAGTGGACCGCGTGCACCCCGTGCACGCGCGCCCGC 660
607 GACACCATCAACGAGAGGCGCCGAGTGGACCGCGTGCACCCCGTGCACGCGCGCCCGC 666
661 ATGCCCCCGGCGCAGATGCGCGAGACCCCGCGGAGCGACATGCGCGGACCAACAGCACC 720
667 ATGCCCCCGGCGCAGATGCGCGAGACCCCGCGGAGCGACATGCGCGGACCAACAGCACC 726
721 CTGCAAGAGAGATGCGCTGTGATGACCAAGACCCCGCATCCCGTGGCGCATCTTAC 780
727 CTGCAAGAGAGATGCGCTGTGATGACCAAGACCCCGCATCCCGTGGCGCATCTTAC 786
781 AAGCGGTGATCATCTGGGCTGAAACAAGATCGTGGCGATGTACAGCCCGGTGAGCATC 840
787 AAGCGGTGATCATCTGGGCTGAAACAAGATCGTGGCGATGTACAGCCCGGTGAGCATC 846
841 CTGACATCAAGCAGAGGCGCCCAAGAGCCCTTCCGCGACTACGTGACCGCTTCTCAAG 900
847 CTGACATCAAGCAGAGGCGCCCAAGAGCCCTTCCGCGACTACGTGACCGCTTCTCAAG 906
901 ACCCTGCGCGCGGAGCAGAGCAACCAAGAGTGAAGAACTGTGATGACCGACACCTGCTG 960
907 ACCCTGCGCGCGGAGCAGAGCAACCAAGAGTGAAGAACTGTGATGACCGACACCTGCTG 966
961 GTGCAAGAGCGCAACCCCGACTGCAAGACCACTCTGCGCGCTCTCGCGCCCGCGCGCAGC 1020
967 GTGCAAGAGCGCAACCCCGACTGCAAGACCACTCTGCGCGCTCTCGCGCCCGCGCGCAGC 1026
1021 CTGAGAGAGATGACCGCTGACAGGCGTGGGCGCGCCCAAGCAAGGCGCGCGT 1080
1027 CTGAGAGAGATGACCGCTGACAGGCGTGGGCGCGCCCAAGCAAGGCGCGCGT 1086
1081 CTGGCCGAGGCGATGAGCAGGCGCAACACCAAGCGTGTATGATCAAGAGCACTTCAAG 1140
1087 CTGGCCGAGGCGATGAGCAGGCGCAACACCAAGCGTGTATGATCAAGAGCACTTCAAG 1146
1141 GGGCCCCCGGCGCATCTCAAGTCTTCAACTGCGGCAAGAGGCGCCACATCGCCCGCAAC 1200
1147 GGGCCCCCGGCGCATCTCAAGTCTTCAACTGCGGCAAGAGGCGCCACATCGCCCGCAAC 1206
1201 TGCCGCGCGCCCGCAAGAGGCGTGTGAAGTGGGCAAGAGGCGCCACATGATGAAG 1260
1207 TGCCGCGCGCCCGCAAGAGGCGTGTGAAGTGGGCAAGAGGCGCCACATGATGAAG 1266
1261 GACTGCACCGAGCGCCAGGCGCACTTCTGGGCAAGATCTGCGCCAGCCCAAGGGCGCGC 1320
1267 GACTGCACCGAGCGCCAGGCGCACTTCTGGGCAAGATCTGCGCGCCAGCCCAAGGGCGCGC 1326
1321 CCGCGCAACTTCTGTGACAGCGCGCGCGGAGCCACCGCGCCCGCGGAGGCTTCCGC 1380
1327 CCGCGCAACTTCTGTGACAGCGCGCGCGGAGCCACCGCGCCCGCGGAGGCTTCCGC 1386
1381 TTGAGAGAGACCAACCCCGCGCGCAAGAGAGAGCAAGAACCGGAGAGACCTTGAAGC 1440
1387 TTGAGAGAGACCAACCCCGCGCGCAAGAGAGAGCAAGAACCGGAGAGACCTTGAAGC 1446
1441 CTGAAGAGCGCTTCTGGCAACGACCCCTGAGCGAGTAA 1479
1447 CTGAAGAGCGCTTCTGGCAACGACCCCTGAGCGAGAAA 1485
```

RESULT 13

```
US-10-190-435-20
; Sequence 20, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Bstrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: P18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 2742
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GagTatRevNef_C
US-10-190-435-20

Query Match          99.8%; Score 1475.8; DB 15; Length 2742;
Best Local Similarity 99.8%; Pred. No. 3.2e-302;
Matches 1477; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 ATGGGCGCGCGCGCCAGCATCTCGCGCGCGCGCAAGCTGAGCGCTGAGGCGCATCCGC 60
DB      7 ATGGGCGCGCGCGCCAGCATCTCGCGCGCGCGCAAGCTGAGCGCTGAGGCGCATCCGC 66
QY      61 CTGCGCGCGCGCGCGCAAGAGTGTACATATGAAGACACTGTGTGGGCGAGCGCGAG 120
DB      67 CTGCGCGCGCGCGCGCAAGAGTGTACATATGAAGACACTGTGTGGGCGAGCGCGAG 126
QY      121 CTGAGAGAGTTCGCTTGAACCCCGCGCTGCTGAGAGACCAAGGCTGCAAGCATC 180
DB      127 CTGAGAGAGTTCGCTTGAACCCCGCGCTGCTGAGAGACCAAGGCTGCAAGCATC 186
QY      181 ATCCGCCAGCTGCACCCCGCGCTGACAGACCGGCGAGGAGCTGAAGAGCTGTTCAC 240
DB      187 ATCCGCCAGCTGCACCCCGCGCTGACAGACCGGCGAGGAGCTGAAGAGCTGTTCAC 246
QY      241 ACCGTGGCCACCTCTACTGCGGTGACAGAGATCGAGGTCCGACACCAAGAGGCC 300
DB      247 ACCGTGGCCACCTCTACTGCGGTGACAGAGATCGAGGTCCGACACCAAGAGGCC 306
QY      301 CTGACCAAGATCGAGAGGAGAGCAACAAGTCCAGAGATCCAGAGGCGGAGGCC 360
DB      307 CTGACCAAGATCGAGAGGAGAGCAACAAGTCCAGAGATCCAGAGGCGGAGGCC 366
QY      361 GCGGCAAGAGGAGGTGAGGCGAGAACTACCCCATGTCAGAACTCTGAGGCGCAGATG 420
DB      367 GCGGCAAGAGGAGGTGAGGCGAGAACTACCCCATGTCAGAACTCTGAGGCGCAGATG 426
QY      421 GTGCACCAAGGCGCATAGCCCCCGCACCTGAAACGCTGTGAAGGTGATCGAGAGAG 480
DB      427 GTGCACCAAGGCGCATAGCCCCCGCACCTGAAACGCTGTGAAGGTGATCGAGAGAG 486
QY      481 GCCTTACAGCCCGGAGGTATCCCATGTTCAACCGCTTGAAGGCGGCGCCACCCCGCAG 540
DB      487 GCCTTACAGCCCGGAGGTATCCCATGTTCAACCGCTTGAAGGCGGCGCCACCCCGCAG 546
QY      541 GACCTGAACAAGATGTTGAACAACCGTGGCGGCCACCAAGCCCGCATGACATGCTGAAG 600
DB      547 GACCTGAACAAGATGTTGAACAACCGTGGCGGCCACCAAGCCCGCATGACATGCTGAAG 606
QY      601 GACACCATCAACGAGAGGCGCGCGAGTGGAGACCGCGTGCACCCCGTGCACGCGCGCCCGC 660
DB      607 GACACCATCAACGAGAGGCGCGCGAGTGGAGACCGCGTGCACCCCGTGCACGCGCGCCCGC 666
QY      661 ATGCCCCCGGCGCAGATGCGCGAGACCCCGCGGAGCGACATGCGCGGACCAACAGCACC 720
```



```
Db ATGCCCCCGCCAGATGCGCGAGCCCCCGCGCAGGCAATGCGCGGCAACCAAGCACC 726
QY CTGCAAGAGAGATGCGCTGTGATGACCAAGACCCCCCATCCCGTGGGCGCATCTAC 780
Db CTGCAAGAGAGATGCGCTGTGATGACCAAGACCCCCCATCCCGTGGGCGCATCTAC 786
QY AAGCGGTGATCATCTGGGCTTGAACAAGATGTGCGGATGTAAGCCCCGTGAGCATC 840
Db AAGCGGTGATCATCTGGGCTTGAACAAGATGTGCGGATGTAAGCCCCGTGAGCATC 846
QY CTGACATCAAGCAGGCGCCCAAGAGACCTTCCGCACTACGTGACCGCTTCTTCAAG 900
Db CTGACATCAAGCAGGCGCCCAAGAGACCTTCCGCACTACGTGACCGCTTCTTCAAG 906
QY ACCCTGCGCGCGAGCAGACACCAAGAGGTGAAGACTGTGATGACCGACACCTGCTG 960
Db ACCCTGCGCGCGAGCAGACACCAAGAGGTGAAGACTGTGATGACCGACACCTGCTG 966
QY GTGCAAGAACCCCAACCCCGACTGCAAGACCATCTGCGCGCTCTCGGCCCGCGCAGC 1020
Db GTGCAAGAACCCCAACCCCGACTGCAAGACCATCTGCGCGCTCTCGGCCCGCGCAGC 1026
QY CTGAGAGAGATGATGACCGCTGCGCAGGCGGTGGCGGCGCCCAAGGCGCGCGTG 1080
Db CTGAGAGAGATGATGACCGCTGCGCAGGCGGTGGCGGCGCCCAAGGCGCGCGTG 1086
QY CTGGCGGAGGTGATGACCGCTGCGCAGGCGGTGGCGGCGCCCAAGGCGCGCGTG 1140
Db CTGGCGGAGGTGATGACCGCTGCGCAGGCGGTGGCGGCGCCCAAGGCGCGCGTG 1146
QY GGGCGCGCGCGCATCTGCAAGTGTCTTCAACTGCGCGCAAGGAGGCGCCATCGCGCGAAC 1200
Db GGGCGCGCGCGCATCTGCAAGTGTCTTCAACTGCGCGCAAGGAGGCGCCATCGCGCGAAC 1206
QY TGCGCGCGCGCGCGCAAGAGGCGCTGTGAGAGTGCGCGCAAGAGGCGCCATGAG 1260
Db TGCGCGCGCGCGCGCAAGAGGCGCTGTGAGAGTGCGCGCAAGAGGCGCCATGAG 1266
QY GACTGCAACGAGCGCGCAAGGCGCACTTCTGCGCAAGATCTGCGCCAGCCCAAGGCGCGC 1320
Db GACTGCAACGAGCGCGCAAGGCGCACTTCTGCGCAAGATCTGCGCCAGCCCAAGGCGCGC 1326
QY CCGCGCACTTCTGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
Db CCGCGCACTTCTGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1386
QY TTGAGAGAGACCAACCCCGCGCGCAAGAGCAGAGAGCAAGGACCGCGAGACCTGACAGC 1440
Db TTGAGAGAGACCAACCCCGCGCGCAAGAGCAGAGAGCAAGGACCGCGAGACCTGACAGC 1446
QY CTGAAGAGCTGTGCGCAAGACCCCTGAGCCAGTAA 1479
Db CTGAAGAGCTGTGCGCAAGACCCCTGAGCCAGTAA 1485
```

## RESULT 14

US-10-190-305A-15

; Sequence 15, Application US/10190305A

; Publication No. US20030198621A1

; GENERAL INFORMATION:

; APPLICANT: ZUR MEGEDE, Jan

; APPLICANT: BARNETT, Susan

; APPLICANT: LIAN, Ying

; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B AND/OR

; TITLE OF INVENTION: TYPE C POLYPEPTIDES, POLYPEPTIDES AND USBS THEREOF

; FILE REFERENCE: 2302-18702 / 18702.002

; CURRENT APPLICATION NUMBER: US/10/190,305A

; CURRENT FILING DATE: 2002-07-05

; NUMBER OF SEQ ID NOS: 93

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 15

; LENGTH: 2742

```
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GagTatRevNef_C
US-10-190-305A-15
```

Query Match 99.8%; Score 1475.8; DB 15; Length 2742;

Best Local Similarity 99.9%; Pred. No. 3.2e-302;

Matches 1477; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 ATGGCGCGCGCGCCAGCATCTGCGCGCGCGCAAGCTGAGCGCTGGAGCGCATCCGC 60
Db 7 ATGGCGCGCGCGCCAGCATCTGCGCGCGCGCAAGCTGAGCGCTGGAGCGCATCCGC 66
QY CTGCGCGCGCGCGCAAGAGGTCTACATGATGAAGCACCTGTGTGGGCCAGCGCGAG 120
Db 67 CTGCGCGCGCGCGCAAGAGGTCTACATGATGAAGCACCTGTGTGGGCCAGCGCGAG 126
QY CTGAGAGAGTTGCGCCCTGAAACCCCGCGCTGTGAGAGACCAAGAGGCGTGCAGCATC 180
Db 127 CTGAGAGAGTTGCGCCCTGAAACCCCGCGCTGTGAGAGACCAAGAGGCGTGCAGCATC 186
QY ATCCGCGAGCTGCAACCCCGCTGACAGACCGCGCAGAGAGCTGAAGAGCCTGTCAAC 240
Db 187 ATCCGCGAGCTGCAACCCCGCTGACAGACCGCGCAGAGAGCTGAAGAGCCTGTCAAC 246
QY ACCGTGCGCACCTGTACTGCGTGCACGAGAGATCGAGGTCCGCAACCAAGAGGCC 300
Db 247 ACCGTGCGCACCTGTACTGCGTGCACGAGAGATCGAGGTCCGCAACCAAGAGGCC 306
QY CTGCACAAGATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
Db 307 CTGCACAAGATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 366
QY GCCGACAAGGCGAAGGTGAGCGCAAGACTACCCCATCGTGCAGAACTGCAGGCGCAGATG 420
Db 367 GCCGACAAGGCGAAGGTGAGCGCAAGACTACCCCATCGTGCAGAACTGCAGGCGCAGATG 426
QY GTGCAACAGGCGCATCAGCCCGCGCACCTGAAGCGCTGGTGAAGGTGATGAGAGAGAG 480
Db 427 GTGCAACAGGCGCATCAGCCCGCGCACCTGAAGCGCTGGTGAAGGTGATGAGAGAGAG 486
QY GCCTTCAAGCGCGGAGGTATCCCATGTTCAACCGCCCTGAGCGAGGCGCCACCCCGCAG 540
Db 487 GCCTTCAAGCGCGGAGGTATCCCATGTTCAACCGCCCTGAGCGAGGCGCCACCCCGCAG 546
QY GACTGGAACAGATGTTGAACACCGTGGCGCGCACAGGCGCGCATGACAGATGTAAG 600
Db 547 GACTGGAACAGATGTTGAACACCGTGGCGCGCACAGGCGCGCATGACAGATGTAAG 606
QY GACACCATCAAGAGAGGCGCGCGCGAGTGGACCGCGTGCACCCCGTGCACGCGCGCCCC 660
Db 607 GACACCATCAAGAGAGGCGCGCGCGAGTGGACCGCGTGCACCCCGTGCACGCGCGCCCC 666
QY ATCGCGCGCGCGCGCAGATGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
Db 667 ATCGCGCGCGCGCGCAGATGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 726
QY CTGCAAGAGAGATGCGCTGTGATGACCAAGACCCCGCGCATGCGCGCGCATCTAC 780
Db 727 CTGCAAGAGAGATGCGCTGTGATGACCAAGACCCCGCGCATGCGCGCGCATCTAC 786
QY AAGCGGTGATCATCTGGGCTTGAACAAGATGTGCGGATGTAAGCCCCGTGAGCATC 840
Db 787 AAGCGGTGATCATCTGGGCTTGAACAAGATGTGCGGATGTAAGCCCCGTGAGCATC 846
QY CTGACATCAAGCAGGCGCGCGCAAGAGCGCTTCCGCGCATACGTGAGACCGCTTCTCAAG 900
Db 847 CTGACATCAAGCAGGCGCGCGCAAGAGCGCTTCCGCGCATACGTGAGACCGCTTCTCAAG 906
QY ACCCTGCGCGCGCGAGCAGAGCACCCAGAGGTGAAGACTGATGACCGACACCTGCTG 960
Db 907 ACCCTGCGCGCGCGAGCAGAGCACCCAGAGGTGAAGACTGATGACCGACACCTGCTG 966
```



QY 1201 TCCCGCGCCCCCGCAAGAGGCTCTGGAAGTGGCAAGAGGGCCACAGATGAG 1260  
|||  
DB 1207 TCCCGCGCCCCCGCAAGAGGCTCTGGAAGTGGCAAGAGGGCCACAGATGAG 1265  
|||  
QY 1261 GACTGCACCGAGCGCCAGGCCAATTCTTGGGCAAGATCTGGCCCAAGGGCGC 1320  
|||  
DB 1267 GACTGCACCGAGCGCCAGGCCAATTCTTGGGCAAGATCTGGCCCAAGGGCGC 1326  
|||  
QY 1321 CCGGCAACTTCTTGCAGAGCGCGCCAGGCCACCGCCCCCGGAGAGCTTCCGC 1380  
|||  
DB 1327 CCGGCAACTTCTTGCAGAGCGCGCCAGGCCACCGCCCCCGGAGAGCTTCCGC 1386  
|||  
QY 1381 TTGAGGAGAGCCACCGCCCGGCAAGCAAGAGCAAGAACCGGCAAGACCTGACCAAC 1440  
|||  
DB 1387 TTGAGGAGAGCCACCGCCCGGCAAGCAAGAGCAAGAACCGGCAAGACCTGACCAAC 1446  
|||  
QY 1441 CTGAAAGAGCTGTTCGGCAAGCAAGACCGCTGAGCCAGTAA 1479  
|||  
DB 1447 CTGAAAGAGCTGTTCGGCAAGCAAGACCGCTGAGCCAGTAA 1485  
|||

Search completed: May 28, 2004, 16:09:19  
Job time : 762.462 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 28, 2004, 07:10:24 ; Search time 685.39 Seconds  
(without alignments)  
9167.164 Million cell updates/sec

Title: US-09-475-704A-3

Perfect score: 1479  
Sequence: 1 atggcgccgcgcgcagcat.....acgacccctgagccagtaa 1479

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_29Jan04:\*

1:	geneseqn1980s:*
2:	geneseqn1990s:*
3:	geneseqn2000s:*
4:	geneseqn2001as:*
5:	geneseqn2001bs:*
6:	geneseqn2002s:*
7:	geneseqn2003as:*
8:	geneseqn2003bs:*
9:	geneseqn2003cs:*
10:	geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1479	100.0	1479	3	AAA51609 HIV synth
2	1479	100.0	1479	6	AAI44548 HIV-1 p55
3	1479	100.0	1479	6	ABL39953 Synthetic
4	1477.4	99.9	3162	9	ADCI3239 DNA of HI
5	1477.4	99.9	3462	9	ADCI3237 DNA of HI
6	1477.4	99.9	4419	7	ACA03523 Synthetic
7	1477.4	99.9	4419	9	ADCI3240 DNA of HI
8	1477.4	99.9	4615	9	ADCI3257 DNA of HI
9	1477.4	99.9	4702	9	ADCI3259 DNA of HI
10	1477.4	99.9	4716	7	ACA03522 Synthetic
11	1477.4	99.9	4716	9	ADCI3238 DNA of HI
12	1475.8	99.8	2742	7	ACA03524 Synthetic
13	1475.8	99.8	2742	9	ADCI3241 DNA of HI
14	1475.8	99.8	3930	9	ADCI3230 DNA of HI
15	1475.8	99.8	3930	9	ADCI3231 DNA of HI
16	1475.8	99.8	3930	9	ADCI3232 DNA of HI
17	1475.8	99.8	5145	7	ACA03521 Synthetic
18	1475.8	99.8	5145	9	ADCI3233 DNA of HI
19	1474.8	99.7	4713	7	ACA03592 Synthetic
20	1474.8	99.7	4713	9	ADCI3280 DNA of HI
21	1473.8	99.6	5184	7	ACA03591 Synthetic
22	1473.8	99.6	5184	9	ADCI3279 DNA of HI
23	1472.4	99.6	2742	7	ACA03590 Synthetic

24	1472.4	99.6	2742	9	ADCI3278	Adci3278 DNA of HI
25	1463	98.9	1479	3	AAA51625	Aaa51625 HIV codon
26	1463	98.9	1479	3	ABL39957	Abi39957 Synthetic
27	1461.4	98.8	1479	6	AAI44552	AaI44552 HIV-1 p55
28	1424.4	96.3	4546	9	ADCI3255	Adci3255 DNA of HI
29	1347.4	91.1	4423	9	ADCI3256	Adci3256 DNA of HI
30	1315.2	88.9	3531	9	ADCI3234	Adci3234 DNA of HI
31	1315.2	88.9	3537	9	ADCI3236	Adci3236 DNA of HI
32	1315.2	88.9	3538	9	ADCI3235	Adci3235 DNA of HI
33	1288.8	87.1	1491	6	ABL40020	AbI40020 Synthetic
34	1288	87.1	1494	6	ABL39972	AbI39972 Synthetic
35	1276.8	86.3	1509	3	AAA51610	Aaa51610 HIV synth
36	1276.8	86.3	1509	6	AAI44549	AaI44549 HIV-1 p55
37	1276.8	86.3	1509	6	ABL39954	AbI39954 Synthetic
38	1262.4	85.4	1509	6	ABL39958	AbI39958 Synthetic
39	1260.8	85.2	1509	6	AAA51626	Aaa51626 HIV codon
40	1260.8	85.2	1509	6	AAI44553	AaI44553 HIV-1 p55
41	1221.2	82.6	9166	6	ABK91616	Abk91616 Modified
42	1206	81.5	9788	6	ABK91622	Abk91622 Modified
43	1205.4	81.5	4288	4	AAH20868	Aah20868 HIV-1 sub
44	1197.8	81.0	1515	3	AAI70412	AaI70412 Synthetic
45	1197.8	81.0	4472	3	AAA70472	Aaa70472 HIV bIcIs

## ALIGNMENTS

## RESULT 1

AAA51609  
ID AAA51609 standard; DNA, 1479 BP.

AC AAA51609;

DT 31-OCT-2000 (first entry)

DE HIV synthetic Gag polynucleotide.

KW Gag; expression cassette; antigenic; type C; HIV; Env; synthetic;

KW DNA immunization; packaging cell line; antigen presentation; ss.

OS Human immunodeficiency virus; type C strain AF110965.

OS Synthetic.

FH Key Location/Qualifiers  
CDS 1..1479

FT /\*tag= a  
FT /product= "Synthetic\_Gag"  
FT /note= "Codon usage pattern was modified and inhibitory  
FT elements (INS) and RRE sites were inactivated resulting  
FT in improved expression"

PN WO200039304-A2.

PD 06-JUL-2000.

PF 30-DEC-1999; 99WO-US031273.

PR 31-DEC-1998; 98US-0114495P.

PR 01-SEP-1999; 99US-0152195P.

PA (CHIR ) CHIRON CORP.

PI Barnett S, Zur Megede J;

DR WPI; 2000-452401/39.

DR P-PSDB; AAY96943.

PT Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV Env  
PT polypeptide and the polypeptide useful for immunizing a mammal especially  
PT human against HIV.

PS Claim 2; Page 92-93; 113pp; English.

CC Expression cassettes comprising a polynucleotide encoding antigenic type  
CC C human immunodeficiency virus (HIV) Gag or Env polypeptides are useful  
CC in DNA immunization, generation of packaging cell lines and production of  
CC Gag- and/or Env-containing proteins. Synthetic Env and Gag expression  
CC cassettes exhibit increased potency for induction of cytotoxic T-  
CC lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1 self-  
CC assemble into non-infectious virus-like particles which are used as a  
CC matrix for the proper presentation of an antigen entrapped or associated  
CC to the immune system of the host

XX  
SQ Sequence 1479 BP; 325 A; 529 C; 463 G; 162 T; 0 U; 0 Other;

Query Match 100.0%; Score 1479; DB 3; Length 1479;

Best Local Similarity 100.0%; Pred. No. 2.3e-185;

Matches 1479; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGGGCGCCCGCCGAGCATCTGCGCGCGGCAAGCTGGAACGCTGGAGCGCATCCGC 60
DB 1 ATGGGCGCCCGCCGAGCATCTGCGCGCGGCAAGCTGGAACGCTGGAGCGCATCCGC 60
QY 61 CTGCGCCCGCGCGCAAGAAGTGTACATGATGAAGCACTGTGTGGCCAGCCGCGAG 120
DB 61 CTGCGCCCGCGCGCAAGAAGTGTACATGATGAAGCACTGTGTGGCCAGCCGCGAG 120
QY 121 CTGGAAGAAGTTCCGCTTGAACCCCGGCTGTCTGAGACCAAGGCGCTGCAAGCATC 180
DB 121 CTGGAAGAAGTTCCGCTTGAACCCCGGCTGTCTGAGACCAAGGCGCTGCAAGCATC 180
QY 181 ATCCGCCAGCTGACCCCGCCCTGCAAGCCGCAAGGAGCTGAAGCCTGTTCAC 240
DB 181 ATCCGCCAGCTGACCCCGCCCTGCAAGCCGCAAGGAGCTGAAGCCTGTTCAC 240
QY 241 ACCGTGGCCACCCCTGTACTGCGTGAAGCAAGAGATCGCGCACCAAGGAGGCC 300
DB 241 ACCGTGGCCACCCCTGTACTGCGTGAAGCAAGAGATCGCGCACCAAGGAGGCC 300
QY 301 CTGGAACAAGATCGAAGAGGAGCAACAAGTGCAGCAAGAATCCAGAGCCGCGGCC 360
DB 301 CTGGAACAAGATCGAAGAGGAGCAACAAGTGCAGCAAGAATCCAGAGCCGCGGCC 360
QY 361 GCCGACAAGGGCAAGGTGAGCCAGAATAACCCATCTGCAAGACCTGCAGGCGCATG 420
DB 361 GCCGACAAGGGCAAGGTGAGCCAGAATAACCCATCTGCAAGACCTGCAGGCGCATG 420
QY 421 GTGCAACCAAGGCGCATCAGCCCGCGCCTTGAACGCTGGTGAAGGTGATCGAAGAAG 480
DB 421 GTGCAACCAAGGCGCATCAGCCCGCGCCTTGAACGCTGGTGAAGGTGATCGAAGAAG 480
QY 481 GCGTTCAAGCCCGGAGGTGATCCCAATTTTCAACGCGCTGAGCGAGGCGCCACCCCGCAG 540
DB 481 GCGTTCAAGCCCGGAGGTGATCCCAATTTTCAACGCGCTGAGCGAGGCGCCACCCCGCAG 540
QY 541 GACCTGAACAAGATGTTGAACAACCGTGGGGCGGCCACAAGCGCCCATGCAATGCTGAAG 600
DB 541 GACCTGAACAAGATGTTGAACAACCGTGGGGCGGCCACAAGCGCCCATGCAATGCTGAAG 600
QY 601 GACACCATCAACGAGAGGCGCGCGAGTGGGACCGCGGTGACACCCGCTGCAGCGCGGCC 660
DB 601 GACACCATCAACGAGAGGCGCGCGAGTGGGACCGCGGTGACACCCGCTGCAGCGCGGCC 660
QY 661 ATCGCCCGCGCGCGAGATGCGGAGCGCGCGCGCGCGCGCATCGCGCGCACCAAGCAACC 720
DB 661 ATCGCCCGCGCGCGAGATGCGGAGCGCGCGCGCGCGCGCATCGCGCGCACCAAGCAACC 720
QY 721 CTGCAAGAGAGATTCGCTGATGACCAACCCCGCATCCCGTGGGCGCAATCTAC 780
DB 721 CTGCAAGAGAGATTCGCTGATGACCAACCCCGCATCCCGTGGGCGCAATCTAC 780
QY 781 AAGCGGTGATCATCTGGGCGCTGAACAAGATCGTGGGATGTACAGCCCGGTGACATC 840
DB 781 AAGCGGTGATCATCTGGGCGCTGAACAAGATCGTGGGATGTACAGCCCGGTGACATC 840
QY 841 CTGGAATCAAGAGGCGCCCAAGAGCCCTTCGCGCACTACGTGACCGCTTCTTCAAG 900
```

```
DB 841 CTGGAATCAAGAGGCGCCCAAGAGCCCTTCGCGCACTACGTGACCGCTTCTTCAAG 900
QY 901 ACCCTGCGCGCGGAGCAGAGCAACCCAGAGGTGAAGACTGATGACCGACCTGCTG 960
DB 901 ACCCTGCGCGCGGAGCAGAGCAACCCAGAGGTGAAGACTGATGACCGACCTGCTG 960
QY 961 GTGCAAGACGCGCAACCCCGCATGCAAGACCATCTGCGCGCTTCGGCCCCGCGCGCAGC 1020
DB 961 GTGCAAGACGCGCAACCCCGCATGCAAGACCATCTGCGCGCTTCGGCCCCGCGCGCAGC 1020
QY 1021 CTGGAAGAGATGATGACCGCGCTGCGCAGGCGCGTGGCGGCCCAAGCAAGGCGCGTG 1080
DB 1021 CTGGAAGAGATGATGACCGCGCTGCGCAGGCGCGTGGCGGCCCAAGCAAGGCGCGTG 1080
QY 1081 CTGCGCCGAGGCGCATGAGCCAGGCGCAACCAAGCGTATGATGCAAGAGCAACTTCAAG 1140
DB 1081 CTGCGCCGAGGCGCATGAGCCAGGCGCAACCAAGCGTATGATGCAAGAGCAACTTCAAG 1140
QY 1141 GCGCCCGCGCGCGCATGCTCAAGTGTCTCACTGCGCAAGAGGCGCATTCGCGCGCAAC 1200
DB 1141 GCGCCCGCGCGCGCATGCTCAAGTGTCTCACTGCGCAAGAGGCGCATTCGCGCGCAAC 1200
QY 1201 TCGCGCGCGCGCGCAAGAGGCGTGTGAAGTGCAGCGCAAGAGGCGCAACGATGAAG 1260
DB 1201 TCGCGCGCGCGCGCAAGAGGCGTGTGAAGTGCAGCGCAAGAGGCGCAACGATGAAG 1260
QY 1261 GACTTCACCGAGCGCGCCAGGCCAATTCTGCGCAAGATCTGCGCGCAAGCGCAAGGCGCGC 1320
DB 1261 GACTTCACCGAGCGCGCCAGGCCAATTCTGCGCGCAAGATCTGCGCGCAAGCGCAAGGCGCGC 1320
QY 1321 CCGCGCAACTTCTCTGCAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 1380
DB 1321 CCGCGCAACTTCTCTGCAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 1380
QY 1381 TTGCAAGAGACCAACCCCGCGCGCAAGAGCAAGAGCAAGAGCAAGAGCGCTGACAGC 1440
DB 1381 TTGCAAGAGACCAACCCCGCGCGCAAGAGCAAGAGCAAGAGCAAGAGCGCTGACAGC 1440
QY 1441 CTGAAGAGCCTGTTTGGCAACGACCCCTGAGCGCATTA 1479
DB 1441 CTGAAGAGCCTGTTTGGCAACGACCCCTGAGCGCATTA 1479
```

#### RESULT 2

AA144548  
ID AA144548 standard; DNA; 1479 BP.

AC AA144548;

XX 29-AUG-2003 (revised)  
DT 08-NOV-2002 (first entry)

XX HIV-1 p55gag polypeptide coding sequence 1.

XX HIV; ds; vaccine; gene; immune response; microparticle;

KW adsorbent surface; poly(alpha-hydroxy acid); polyhydroxy butyric acid;

KW polycaprolactone; polyorthoester; polycyanoacrylate; detergent;

KW submicron emulsion; viral infection; bacterial infection;

XX parasitic infection; HIV-1 p55gag polypeptide.

OS Human immunodeficiency virus 1.

XX WO200226209-A2.

XX 04-APR-2002.

XX 28-SEP-2001; 2001WO-US030540.

XX 28-SEP-2000; 2000US-0236105P.

XX 30-AUG-2001; 2001US-0315905P.

XX (CHIR ) CHIRON CORP.



XX O'hagan D, Otten G, Donnelly JU, Polo JM, Barnett S, Singh M;  
PI Ulmer J, Dubensky TW;  
XX  
DR WPI; 2002-519084/55.  
XX  
PT A microparticle to which a biologically active macromolecule is adsorbed,  
PT for use as a vaccine composition to treat viral, bacterial or parasitic  
PT infections, comprises a polymer microparticle, a detergent and a  
PT submicron emulsion.  
XX  
PS Claim 72; Fig 1; 100pp; English.  
XX  
CC The invention relates to a method of raising an immune response in a host  
CC animal. The method of the invention comprises administering a  
CC microparticle that has an adsorbent surface to which a first biologically  
CC active macromolecule (e.g. a nucleic acid) has been adsorbed. The  
CC microparticle comprises a polymer microparticle of poly(alpha-hydroxy  
CC acid), a polyhydroxy butyric acid, a polycaprolactone, a polyorthoester,  
CC a polycyanoacrylate, a detergent, and submicron emulsion. The method/  
CC microparticle of the invention is useful for immunising a host animal  
CC against viral, bacterial or parasitic infections. The present DNA  
CC sequence encodes a HIV-1 p55gag polypeptide. (Updated on 29-AUG-2003 to  
CC standardise OS field)  
XX  
SQ Sequence 1479 BP; 325 A; 529 C; 463 G; 162 T; 0 U; 0 Other;

Query Match 100.0%; Score 1479; DB 6; Length 1479;  
Best Local Similarity 100.0%; Pred. No. 2.3e-185;  
Matches 1479; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGGGGCGCGCGCCAGCATCTCTCGCGCGCGCGCAAGCTGAGCCCTGGAGCGCATCCGC 60  
DB 1 ATGGGGCGCGCGCGCCAGCATCTCTCGCGCGCGCGCAAGCTGAGCCCTGGAGCGCATCCGC 60  
OY 61 CTGCGCGCGCGCGCGCGCAAGAGTCTACATGATGAAGCACTGTGTGTGGCCAGCGCGAG 120  
DB 61 CTGCGCGCGCGCGCGCGCAAGAGTCTACATGATGAAGCACTGTGTGTGGCCAGCGCGAG 120  
OY 121 CTGAGAGAGTTGCGCCCTGAACCCCGCGCTGTGAGAGACCAAGGGCTGCAAGCATC 180  
DB 121 CTGAGAGAGTTGCGCCCTGAACCCCGCGCTGTGAGAGACCAAGGGCTGCAAGCATC 180  
OY 181 ATCCGCCAGCTGACACCCCGCGCTGACAGCCGCGAGCGAGCTGAAGGCTTTCAC 240  
DB 181 ATCCGCCAGCTGACACCCCGCGCTGACAGCCGCGAGCGAGCTGAAGGCTTTCAC 240  
OY 241 ACCGTGGCCACCTCTGTACTGTGCTGACAGAGAGATGAGGTCCGCAACCAAGAGGCC 300  
DB 241 ACCGTGGCCACCTCTGTACTGTGCTGACAGAGAGATGAGGTCCGCAACCAAGAGGCC 300  
OY 301 CTGAGCAAGATCGAGAGAGAGAGAGAGAGATGAGGTCCGCAACCAAGAGGCC 360  
DB 301 CTGAGCAAGATCGAGAGAGAGAGAGAGAGATGAGGTCCGCAACCAAGAGGCC 360  
OY 361 GCCGACAGAGGCAAGGTGAGGCAAGACTACCCCATCTGTCAGAACTGCAAGGCCAGATG 420  
DB 361 GCCGACAGAGGCAAGGTGAGGCAAGACTACCCCATCTGTCAGAACTGCAAGGCCAGATG 420  
OY 421 GTGCACCAAGGCGCATCAGCCCCCGCACCTTGAACGCTTGGGTGAAGGTGATCGAGAGAG 480  
DB 421 GTGCACCAAGGCGCATCAGCCCCCGCACCTTGAACGCTTGGGTGAAGGTGATCGAGAGAG 480  
OY 481 GCCTTCAGGCTCCGAGGTGATCCCATGTTCAACCGCCCTGAGGAGGGCGCCACCCCCAG 540  
DB 481 GCCTTCAGGCTCCGAGGTGATCCCATGTTCAACCGCCCTGAGGAGGGCGCCACCCCCAG 540  
OY 541 GACCTGAACACGATGTTGAACACCGTGGCGGCGCACAGGCGCCCATGCAAGTCTGAAG 600  
DB 541 GACCTGAACACGATGTTGAACACCGTGGCGGCGCACAGGCGCCCATGCAAGTCTGAAG 600  
OY 601 GACACCATCAACGAGAGGCGCGCGAGTGGAGCCGCGTGCACCCCGTGCAGCGCGCGCCC 660  
DB 601 GACACCATCAACGAGAGGCGCGCGAGTGGAGCCGCGTGCACCCCGTGCAGCGCGCGCCC 660

DB 601 GACACCATCAACGAGAGGCGCGCGAGTGGAGCCGCGTGCACCCCGTGCAGCGCGCGCCC 660  
OY 661 ATGCGCGCGCGCGCGAGATGCGGAGCGCGCGCGCGAGGACATGCGCGGCAACACGACACC 720  
DB 661 ATGCGCGCGCGCGCGAGATGCGGAGCGCGCGCGCGAGGACATGCGCGGCAACACGACACC 720  
OY 721 CTGACAGAGCAGATCGCTGTGATGACCAAGAACCCCTCATCCCGTGGGGGACATCTAC 780  
DB 721 CTGACAGAGCAGATCGCTGTGATGACCAAGAACCCCTCATCCCGTGGGGGACATCTAC 780  
OY 781 AAGCGGTGATCATCTTGGGCTTGAACAAGATCTGCGGATGTACAGCCCGTGAGATC 840  
DB 781 AAGCGGTGATCATCTTGGGCTTGAACAAGATCTGCGGATGTACAGCCCGTGAGATC 840  
OY 841 CTGACATCAAGAGGCG 900  
DB 841 CTGACATCAAGAGGCG 900  
OY 901 ACCCTGCG 960  
DB 901 ACCCTGCG 960  
OY 961 GTGACAGACGCGCAACCCCGACTGCAAGACCATCTGCGGCTTCTGCGCGCGCGCGCG 1020  
DB 961 GTGACAGACGCGCAACCCCGACTGCAAGACCATCTGCGGCTTCTGCGCGCGCGCGCG 1020  
OY 1021 CTGAGAGAGATGATGACCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1080  
DB 1021 CTGAGAGAGATGATGACCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1080  
OY 1081 CTGCGCGAGGCGATGAGCCAGGCGCAACACCAAGCTGATGATGAGAGAGCAACTTCAAG 1140  
DB 1081 CTGCGCGAGGCGATGAGCCAGGCGCAACACCAAGCTGATGATGAGAGAGCAACTTCAAG 1140  
OY 1141 GCG 1200  
DB 1141 GCG 1200  
OY 1201 TGCG 1260  
DB 1201 TGCG 1260  
OY 1261 GACTGCACCGAGCG 1320  
DB 1261 GACTGCACCGAGCG 1320  
OY 1321 CCGCGCAACTTCTGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380  
DB 1321 CCGCGCAACTTCTGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380  
OY 1381 TTGAGAGAGACCAACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1440  
DB 1381 TTGAGAGAGACCAACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1440  
OY 1441 CTGAAGAGGCTGTTCGCAAGCAAGCCCTGAGCCAGTAA 1479  
DB 1441 CTGAAGAGGCTGTTCGCAAGCAAGCCCTGAGCCAGTAA 1479

RESULT 3  
ABL39953  
ID ABL39953 standard; DNA; 1479 BP.  
XX  
AC ABL39953;  
XX  
DT 15-MAY-2002 (first entry)  
XX  
DE Synthetic Gag polynucleotide sequence SEQ ID NO:3.  
XX  
KW Human immunodeficiency virus type C; antigenic HIV type C protein;  
KW immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;  
KW immunostimulant; gene therapy; gene; ds.  
XX

OS Human immunodeficiency virus; type C.  
OS Synthetic.

XX WO200204493-A2.

XX 17-JAN-2002.

XX 05-JUL-2001; 2001WO-US021241.

XX 05-JUL-2000; 2000US-00610313.

XX (CHIR ) CHIRON CORP.  
XX (UYST-) UNIV STELLENBOSCH.

XX Zur Megele J, Barnett SW, Engelbrecht S, Van Rensburg BJ;

XX MPI; 2002-154920/20.

XX New polynucleotides encoding antigenic HIV Type C polypeptides, useful in  
XX applications including DNA immunization or generation of packaging cell  
XX lines, particularly in gene therapy.

XX Example 1; Fig 1; 233pp; English.

XX The present invention describes expression cassettes comprising a  
XX polynucleotide sequence encoding a polypeptide comprising immunogenic HIV  
XX type C polypeptides. The expression cassettes comprise any of the HIV  
XX type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or Nef  
XX (I). (I) have immunostimulant activity and can be used in gene therapy.  
XX The HIV type C polynucleotides are useful in applications including DNA  
XX immunisation, generation of packaging cell lines, and production of HIV  
XX type C proteins. The polynucleotides are particularly useful in gene  
XX therapy and DNA immunisation applications. ABL3942 to ABL40054 and  
XX ABB06204 to ABB06215 represent sequences used in the exemplification of  
XX the present invention

XX Sequence 1479 BP; 325 A; 529 C; 463 G; 162 T; 0 U; 0 Other;

XX Query Match 100.0%; Score 1479; DB 6; Length 1479;

XX Best Local Similarity 100.0%; Pred. No. 2.3e-185;

XX Matches 1479; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 1 ATGGGGCCCCCGCCAGCATCTGCGCGCGCGGCAAGCTGACCGCTGGAGCGCATCCGC 60

XX 1 ATGGGGCCCCCGCCAGCATCTGCGCGCGCGGCAAGCTGACCGCTGGAGCGCATCCGC 60

XX 61 CTGCGCCCCCGCGGCAAGAGTCTCATGATGAGCACTGTGTGGGCGGCGGAG 120

XX 61 CTGCGCCCCCGCGGCAAGAGTCTCATGATGAGCACTGTGTGGGCGGCGGAG 120

XX 121 CTGGAAGAGTTCCTCTGAAACCCCGGCTGTGAGACCAAGGCGCTGCAAGAGATC 180

XX 121 CTGGAAGAGTTCCTCTGAAACCCCGGCTGTGAGACCAAGGCGCTGCAAGAGATC 180

XX 181 ATCCGCCAGCTGACCCCGGCTGTGAGACCAAGGCGCTGCAAGAGATC 240

XX 181 ATCCGCCAGCTGACCCCGGCTGTGAGACCAAGGCGCTGCAAGAGATC 240

XX 241 ACCGTGGCCACCTCTGTACTGCGTGCAGAGAGATCGAGTCCGCGACACCAAGAGGCC 300

XX 241 ACCGTGGCCACCTCTGTACTGCGTGCAGAGAGATCGAGTCCGCGACACCAAGAGGCC 300

XX 301 CTGGAACAAGATCGAGAGAGAGCAACAAGTGCAGCAAGAATCGAGAGGCCGAGGCC 360

XX 301 CTGGAACAAGATCGAGAGAGAGCAACAAGTGCAGCAAGAATCGAGAGGCCGAGGCC 360

XX 361 GCCGACAAGGCAAGTGAAGCCAGAACTAACCCCATCTGTGAGAACCTGCAAGGCGCATG 420

XX 361 GCCGACAAGGCAAGTGAAGCCAGAACTAACCCCATCTGTGAGAACCTGCAAGGCGCATG 420

XX 421 GTGCAACCAAGGCATCAAGCCCGGCACTTGAACGCTGGGTGAAGGTGATCGAGAGAAAG 480

XX 421 GTGCAACCAAGGCATCAAGCCCGGCACTTGAACGCTGGGTGAAGGTGATCGAGAGAAAG 480

QY 481 GCCTTCAGCCCCGAGGTGATCCCATGTTTCAACCGCCCTGAGCGAGGGCGCCACCCCCAG 540

DB 481 GCCTTCAGCCCCGAGGTGATCCCATGTTTCAACCGCCCTGAGCGAGGGCGCCACCCCCAG 540

QY 541 GACCTGAACAGATGTTGAACACCGTGGGCGGCAACGAGCGCCATGCAAGTCTGAAG 600

DB 541 GACCTGAACAGATGTTGAACACCGTGGGCGGCAACGAGCGCCATGCAAGTCTGAAG 600

QY 601 GACACCATCAACGAGAGAGGCGCGGAGTGGGACCGGTCACCCCGTGACAGCGCGGCC 660

DB 601 GACACCATCAACGAGAGAGGCGCGGAGTGGGACCGGTCACCCCGTGACAGCGCGGCC 660

QY 661 ATCGCCCCCGGCAAGATGCGCGAGCGCCCGGCGAGCGACATCGCCGCGCACACGACCC 720

DB 661 ATCGCCCCCGGCAAGATGCGCGAGCGCCCGGCGAGCGACATCGCCGCGCACACGACCC 720

QY 721 CTGCAAGAGAGATCGCTGTGATGACCAAGACCCCGCATCCCGTGAGGCGCATCTAC 780

DB 721 CTGCAAGAGAGATCGCTGTGATGACCAAGACCCCGCATCCCGTGAGGCGCATCTAC 780

QY 781 AAGCGGTGATCATCTGCGCTGAACAAGATCGTGGGATTAACAGCCCGTGACATC 840

DB 781 AAGCGGTGATCATCTGCGCTGAACAAGATCGTGGGATTAACAGCCCGTGACATC 840

QY 841 CTGGAATCAAGCAGAGGCGCCCAAGAGAGCCCTTCGCGACTAGTGGACCGCTTCAAG 900

DB 841 CTGGAATCAAGCAGAGGCGCCCAAGAGAGCCCTTCGCGACTAGTGGACCGCTTCAAG 900

QY 901 ACCCTGCGCGCGAGCAGAGACCAAGAGTGAAGAACTGATGACCGACCTGCTG 960

DB 901 ACCCTGCGCGCGAGCAGAGACCAAGAGTGAAGAACTGATGACCGACCTGCTG 960

QY 961 GTGCAAGACCGCAACCCCGACTGCAAGACCACTCTGCGCGCTCTCGGCCCGGCGCAGC 1020

DB 961 GTGCAAGACCGCAACCCCGACTGCAAGACCACTCTGCGCGCTCTCGGCCCGGCGCAGC 1020

QY 1021 CTGGAAGAGATGATGACCGCTGTGCAAGGCGGCGGCGGCGGCGGCGGCGGCGG 1080

DB 1021 CTGGAAGAGATGATGACCGCTGTGCAAGGCGGCGGCGGCGGCGGCGGCGGCGG 1080

QY 1081 CTGGCCGAGGCGGATGAGCCAGGCGCAACCAAGCGTGTATGACAGAGCAACTTCAAG 1140

DB 1081 CTGGCCGAGGCGGATGAGCCAGGCGCAACCAAGCGTGTATGACAGAGCAACTTCAAG 1140

QY 1141 GGGCCCCCGGCGCATCTGTAAGTCTTCAACTGCGGCAAGAGGCGCATCGCGCGAAC 1200

DB 1141 GGGCCCCCGGCGCATCTGTAAGTCTTCAACTGCGGCAAGAGGCGCATCGCGCGAAC 1200

QY 1201 TGCCTGCGCGCGCGCAAGAGGCGTCTGGAAGTCCGCAAGAGGCGCCACCAAGTGAAG 1260

DB 1201 TGCCTGCGCGCGCGCAAGAGGCGTCTGGAAGTCCGCAAGAGGCGCCACCAAGTGAAG 1260

QY 1261 GACTGCACCGAGCGCCAGGCGCAACTTCTGCGGCAAGATCTGCGGCAAGAGGCGCGC 1320

DB 1261 GACTGCACCGAGCGCCAGGCGCAACTTCTGCGGCAAGATCTGCGGCAAGAGGCGCGC 1320

QY 1321 CCGGCGCAACTTCTGCAAGAGCGCGCCGAGCCACCGCCCGCGCGCGGAGAGCTTCCGC 1380

DB 1321 CCGGCGCAACTTCTGCAAGAGCGCGCCGAGCCACCGCCCGCGCGGAGAGCTTCCGC 1380

QY 1381 TTGAGAGAGACCAACCCCGGCGCAAGAGCAAGAGCAAGAGCAAGAGCAAGAGCAAG 1440

DB 1381 TTGAGAGAGACCAACCCCGGCGCAAGAGCAAGAGCAAGAGCAAGAGCAAGAGCAAG 1440

QY 1441 CTGAAGAGCTGTTCGCAACGACCCCTGAGCGCATTA 1479

DB 1441 CTGAAGAGCTGTTCGCAACGACCCCTGAGCGCATTA 1479

RESULT 4  
ADCL3239  
ID ADCL3239 standard; DNA; 3162 BP.

XX AC ADL3239;  
XX DT 16-DEC-2003 (first entry)  
XX DE DNA of HIV construct GagRmut\_C SBQ ID NO 18.  
XX expression cassette; HIV Gag; Env; Int; Nef; p15RnaseH; Pol; Tat; Prot;  
XX Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; ds.  
XX OS Human immunodeficiency virus.  
XX PN WO2003004620-A2.  
XX PD 16-JAN-2003.  
XX PF 05-JUL-2002; 2002WO-US021420.  
XX PR 05-JUL-2001; 2001US-0303192P.  
XX PR 31-AUG-2001; 2001US-0316860P.  
XX PR 16-JAN-2002; 2002US-0349871P.  
XX PA (CHIR ) CHIRON CORP.  
XX PA (UYST-) UNIV STELLENBOSCH.  
XX PI Zur Megede J, Barnett SW, Lian Y, Engelbrecht S, Van Rensburg BJ;  
XX DR WPI; 2003-221593/21.  
XX PT New expression cassette comprising a polynucleotide sequence encoding a  
XX PT polypeptide including an HIV Gag, Env, Int, Nef, p15RnaseH, Pol, Tat,  
XX PT Prot, or Rev polypeptide, useful for immunization, or generating  
XX PT packaging cell lines.  
XX PS Disclosure; Fig 15; 301pp; English.  
XX CC The invention relates to a novel expression cassette comprising a  
XX CC polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,  
XX CC Int, Nef, p15RnaseH, Pol, Tat, Prot, or Rev polypeptide. The novel  
XX CC expression cassette can be used to treat HIV type C by gene therapy or  
XX CC used in the development of a vaccine. The gene delivery vector is  
XX CC administered intramuscularly, intracutaneously, intranasally,  
XX CC subcutaneously, intradermally, transdermally, intravaginally,  
XX CC intrarectally, orally or intravenously. The expression cassette is useful  
XX CC for immunisation, generating packaging cell lines and producing HIV  
XX CC polypeptides. This polynucleotide sequence represents the DNA of an HIV  
XX CC Type C related sequence of the invention.  
SQ Sequence 3162 BP; 729 A; 1095 C; 975 G; 363 T; 0 U; 0 Other;  
Query Match 99.9%; Score 1477.4; DB 9; Length 3162;  
Best Local Similarity 99.9%; Pred. No. 3.3e-185;  
Matches 1478; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 ATGGGGCGCGCGCAGATCTGCGCGCGCGCAAGCTGAGCGCTGGGAGCGCATCCGC 60  
DB 7 ATGGGGCGCGCGCAGATCTGCGCGCGCGCAAGCTGAGCGCTGGGAGCGCATCCGC 66  
QY 61 CTGCGCGCGCGCGCAGAGAGTGCTACATGATGAAGCACTGTGTGGGCCAGCGCGAG 120  
DB 67 CTGCGCGCGCGCGCAGAGAGTGCTACATGATGAAGCACTGTGTGGGCCAGCGCGAG 126  
QY 121 CTGGAAGAAGTTGCGCTGAACCCCGCGCTGTGAGACGAGCGAGGCTGCAAGCAGATC 180  
DB 127 CTGGAAGAAGTTGCGCTGAACCCCGCGCTGTGAGACGAGCGAGGCTGCAAGCAGATC 186  
QY 181 ATCGCGCAGCTGACACCGCGCGCTGCAAGCGGCGAGCGAGAGCTGAAGAGCTGTTCAAC 240  
DB 187 ATCGCGCAGCTGACACCGCGCGCTGCAAGCGGCGAGCGAGAGCTGAAGAGCTGTTCAAC 246  
QY 241 ACCGTGGCCACCTGTACTGCGTGAAGAGAGATGAGGTCCGCGACCAAGAGGCGC 300  
DB 247 ACCGTGGCCACCTGTACTGCGTGAAGAGAGATGAGGTCCGCGACCAAGAGGCGC 306

QY 301 CTGACAAAGATCGAGAGAGCAGAAACAAGTCCAGCAGAAAGATCCAGAGCGCGAGCC 360  
DB 307 CTGACAAAGATCGAGAGAGCAGAAACAAGTCCAGCAGAAAGATCCAGAGCGCGAGCC 366  
QY 361 GCCGACAAAGGCAAGGTGAGCCAGAACTAACCCCATCTGTGAGAACTTGACAGGCGCAGATG 420  
DB 367 GCCGACAAAGGCAAGGTGAGCCAGAACTAACCCCATCTGTGAGAACTTGACAGGCGCAGATG 426  
QY 421 GTGCACCAAGGCATCAAGCCCGCGCACCTGAACGCTGGGTGAAGGTGATCGAGAGAAAG 480  
DB 427 GTGCACCAAGGCATCAAGCCCGCGCACCTGAACGCTGGGTGAAGGTGATCGAGAGAAAG 486  
QY 481 GCCTTCAGCCCGAGGTGATCCCATGTTCACCGCCCTGAGCGAGGCGGCCACCCCCAG 540  
DB 487 GCCTTCAGCCCGAGGTGATCCCATGTTCACCGCCCTGAGCGAGGCGGCCACCCCCAG 546  
QY 541 GACCTGAACAAGATGTTGAACACCGTGGCGCGCCACACAGCGCCCATGCAATGCTGAAG 600  
DB 547 GACCTGAACAAGATGTTGAACACCGTGGCGCGCCACACAGCGCCCATGCAATGCTGAAG 606  
QY 601 GACACCATCAACGAGAGGCGCGCGAGTGGGACCGCGTGCAACCCGTGCAACCGCGGCC 660  
DB 607 GACACCATCAACGAGAGGCGCGCGAGTGGGACCGCGTGCAACCCGTGCAACCGCGGCC 666  
QY 661 ATGCCCCCGCGCAGATGCGCGAGCGCGCGCGCAGCGACATCGCGCGCACACAGCACC 720  
DB 667 ATGCCCCCGCGCAGATGCGCGAGCGCGCGCGCAGCGACATCGCGCGCACACAGCACC 726  
QY 721 CTGAGAGCAGATGCGCTGATGACCAACACCCCATCCCGTGGCGCAGATCTAC 780  
DB 727 CTGAGAGCAGATGCGCTGATGACCAACACCCCATCCCGTGGCGCAGATCTAC 786  
QY 781 AAGCGTGTATCATCTGGGCTGAACAAGATGTCGCGATGTACAGCCCGTGAGCATC 840  
DB 787 AAGCGTGTATCATCTGGGCTGAACAAGATGTCGCGATGTACAGCCCGTGAGCATC 846  
QY 841 CTGACATCAAGCAGAGGCGCGCAAGAGGCTTCCGCGACTACGTGACCGCTTCTCAAG 900  
DB 847 CTGACATCAAGCAGAGGCGCGCAAGAGGCTTCCGCGACTACGTGACCGCTTCTCAAG 906  
QY 901 ACCCTGCGCGCGCAGCAGACACCCAGAGGTGAAGACTGATGACCGACACCTGCTG 960  
DB 907 ACCCTGCGCGCGCAGCAGACACCCAGAGGTGAAGACTGATGACCGACACCTGCTG 966  
QY 961 GTGCAGAACGCCAACCCCGACTGCAAGACATCTCGCGCTCTCGCGCGCGCGCAGC 1020  
DB 967 GTGCAGAACGCCAACCCCGACTGCAAGACATCTCGCGCTCTCGCGCGCGCGCAGC 1026  
QY 1021 CTGAGAGATGATGACCGCTGCGCAGGCGGTGGCGCGCCAGCCACAGCGCCCGCTG 1080  
DB 1027 CTGAGAGATGATGACCGCTGCGCAGGCGGTGGCGCGCCAGCCACAGCGCCCGCTG 1086  
QY 1081 CTGCGCGAGCGATGAGCCAGGCGCAACACACGCTGATGATGCAAGAGCACTTCAAG 1140  
DB 1087 CTGCGCGAGCGATGAGCCAGGCGCAACACACGCTGATGATGCAAGAGCACTTCAAG 1146  
QY 1141 GCGCGCGCGCGCATGCTGAAGTCTTCAACTGCGCGCAAGAGGCGCAACTTCCGCAAC 1200  
DB 1147 GCGCGCGCGCGCATGCTGAAGTCTTCAACTGCGCGCAAGAGGCGCAACTTCCGCAAC 1206  
QY 1201 TGC CGCGCGCGCGCAAGAGGCTGCTGAAGTGC CGGCAAGAGGCGCAACAGATGAAG 1260  
DB 1207 TGC CGCGCGCGCGCAAGAGGCTGCTGAAGTGC CGGCAAGAGGCGCAACAGATGAAG 1266  
QY 1261 GACTGCAACGAGCGCGCAGCCCACTTCTGCGCGCAAGATCTGGCCCAAGGCGCGC 1320  
DB 1267 GACTGCAACGAGCGCGCAGCCCACTTCTGCGCGCAAGATCTGGCCCAAGGCGCGC 1326  
QY 1321 CCGCGCAACTTCTGCAAGCGCGCGCGCAGCCCAACCGCGCGCGCGCGCAGAGCTTCCGC 1380  
DB 1327 CCGCGCAACTTCTGCAAGCGCGCGCGCAGCCCAACCGCGCGCGCGCGCAGAGCTTCCGC 1386

QY 1381 TTGAGAGACACACCCCGGCAAGAGCAGAGACCGGAGACCTGACACG 1440  
DB 1387 TTGAGAGACACACCCCGGCAAGAGCAGAGACCGGAGACCTGACACG 1446  
QY 1441 CTGAAGAGCCTGTGCGCAAGACACCCCTGAGCCAGTAA 1479  
DB 1447 CTGAAGAGCCTGTGCGCAAGACACCCCTGAGCCAGAAA 1485

RESULT 5  
ADCl3237  
ID ADCl3237 standard; DNA; 3462 BP.  
XX ADCl3237;  
XX 18-DEC-2003 (first entry)  
DB DNA of HIV construct GagProtiArlmut\_C SEQ ID NO 16.  
XX expression cassette; HIV Gag; Env; Int; Nef; p15RaseH; Pol; Tat; Prot;  
KW Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; de-  
XX Human Immunodeficiency virus.  
XX WO2003004620-A2.  
XX PN 16-JAN-2003.  
XX PD 05-JUL-2002; 2002WO-US021420.  
XX PR 05-JUL-2001; 2001US-0303192P.  
XX PR 31-AUG-2001; 2001US-0316860P.  
XX PR 16-JAN-2002; 2002US-0349871P.  
XX PA (CHIR ) CHIRON CORP.  
XX PA (UYST-) UNIV STIELEBOSCH.  
XX PI Zur Megede J, Barnett SW, Lian Y, Engelbrecht S, Van Rensburg BJ;  
XX WPI; 2003-221593/21.  
XX New expression cassette comprising a polynucleotide sequence encoding a  
PT polypeptide including an HIV Gag, Env, Int, Nef, p15RaseH, Pol, Tat,  
PT Prot, or Rev polypeptide, useful for immunization, or generating  
PT packaging cell lines.  
XX Disclosure; Fig 13; 301pp; English.  
XX The invention relates to a novel expression cassette comprising a  
CC polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,  
CC Int, Nef, p15RaseH, Pol, Tat, Prot, or Rev polypeptide. The novel  
CC expression cassette can be used to treat HIV type C by gene therapy or  
CC used in the development of a vaccine. The gene delivery vector is  
CC administered intramuscularly, intramuscularly, intranasally,  
CC subcutaneously, intradermally, transdermally, intravaginally,  
CC intrarectally, orally or intravenously. The expression cassette is useful  
CC for immunisation, generating packaging cell lines and producing HIV  
CC polypeptides. This polynucleotide sequence represents the DNA of an HIV  
CC Type C related sequence of the invention.  
XX  
SQ Sequence 3462 BP; 790 A; 1198 C; 1068 G; 406 T; 0 U; 0 Other;

Query Match 99.9%; Score 1477.4; DB 9; Length 3462;  
Best Local Similarity 99.9%; Pred. No. 3.2e-185;  
Matches 1478; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGCGCGCGCAGCATCTGCGCGCGCGCAAGCTGAGCGCTGGAGCGCATCCGC 60  
DB 7 ATGGGCGCGCGCGCAGCATCTGCGCGCGCGCAAGCTGAGCGCTGGAGCGCATCCGC 66  
QY 61 CTGGCGCGCGCGCGCAAGAGTGCTACATGATGAAGCACTGTGTGTGGCGCGCGGAG 120  
DB 67 CTGGCGCGCGCGCGCAAGAGTGCTACATGATGAAGCACTGTGTGTGGCGCGCGGAG 126

QY 121 CTGAGAGAGTTGCGCTTGAACCCCGGCTGCTGAGACCGAGCGGCTGCAAGCAGATC 180  
DB 127 CTGAGAGAGTTGCGCTTGAACCCCGGCTGCTGAGACCGAGCGGCTGCAAGCAGATC 186  
QY 181 ATCCGCGAGCTGACACCCCGGCTGACAGACCGGCGAGGAGCTGAAGAGCTGTTCAC 240  
DB 187 ATCCGCGAGCTGACACCCCGGCTGACAGACCGGCGAGGAGCTGAAGAGCTGTTCAC 246  
QY 241 ACCGTGCGCAACCTGTACTGCTGTCAGCAGAGAGATCGAGGTCCGCAACCAAGAGGCG 300  
DB 247 ACCGTGCGCAACCTGTACTGCTGTCAGCAGAGAGATCGAGGTCCGCAACCAAGAGGCG 306  
QY 301 CTGAGCAAGATCGAGAGAGCAGAAACAAGTCCAGCAAGATCCAGAGCGCGAGGCC 360  
DB 307 CTGAGCAAGATCGAGAGAGCAGAAACAAGTCCAGCAAGATCCAGAGCGCGAGGCC 366  
QY 361 GCGGACAAAGGCAAGTGAGCCCAAGACTACCCCATCTGCGAAGACCTGACAGGCGCAGATG 420  
DB 367 GCGGACAAAGGCAAGTGAGCCCAAGACTACCCCATCTGCGAAGACCTGACAGGCGCAGATG 426  
QY 421 GTGCACCAAGGCGCATCAGCCCGGCAACCTGAAAGCTGGGTGAAGGTATGAGAGAAAG 480  
DB 427 GTGCACCAAGGCGCATCAGCCCGGCAACCTGAAAGCTGGGTGAAGGTATGAGAGAAAG 486  
QY 481 GCCTTACAGCCCGGAGTGATTCCTCCATGTTCAACCGGCTGAGCGAGGCGGCCACCCCGAG 540  
DB 487 GCCTTACAGCCCGGAGTGATTCCTCCATGTTCAACCGGCTGAGCGAGGCGGCCACCCCGAG 546  
QY 541 GACCTGAACAGATGTTGAACAACGTGGGCGGCCACAGGCGCCATGACAGATGCTGAAG 600  
DB 547 GACCTGAACAGATGTTGAACAACGTGGGCGGCCACAGGCGCCATGACAGATGCTGAAG 606  
QY 601 GACACCATCAACGAGAGAGCGCGCGAGTGAGACCGGCGTGCACCCCGTGACGCGCGGCC 660  
DB 607 GACACCATCAACGAGAGAGCGCGCGAGTGAGACCGGCGTGCACCCCGTGACGCGCGGCC 666  
QY 661 ATCGCCCCCGGCGCAGATGCGCGAGAGCCCGGCGGCGAGCAGATCGCGCGGACCAAGCAGCC 720  
DB 667 ATCGCCCCCGGCGCAGATGCGCGAGAGCCCGGCGGCGAGCAGATCGCGCGGACCAAGCAGCC 726  
QY 721 CTGAGAGAGAGATCGCTGATGATGACCAAGACACCCCGCATTCCTGGGCGCATCTAC 780  
DB 727 CTGAGAGAGAGATCGCTGATGATGACCAAGACACCCCGCATTCCTGGGCGCATCTAC 786  
QY 781 AAGCGGTGATCATCTGCGGCTGAACAAGATCTGCGAGATGTAACAGCCCGGTGAGCATC 840  
DB 787 AAGCGGTGATCATCTGCGGCTGAACAAGATCTGCGAGATGTAACAGCCCGGTGAGCATC 846  
QY 841 CTGAGCATCAAGCAGAGGCGGCCCAAGAGGCGCTTCGCGACTACGTGAGCGCTTCTTAAG 900  
DB 847 CTGAGCATCAAGCAGAGGCGGCCCAAGAGGCGCTTCGCGACTACGTGAGCGCTTCTTAAG 906  
QY 901 ACCCTGCGCGCGGAGCAGAGCAGCAGAGAGGTGAAGAACTGATGACCGACACCTGCTG 960  
DB 907 ACCCTGCGCGCGGAGCAGAGCAGCAGAGAGGTGAAGAACTGATGACCGACACCTGCTG 966  
QY 961 GTGCAAGACCGCAACCCCGACTGCAAGACCATCTGCGCGCTCTGCGCGCGCGCGCAGC 1020  
DB 967 GTGCAAGACCGCAACCCCGACTGCAAGACCATCTGCGCGCTCTGCGCGCGCGCGCAGC 1026  
QY 1021 CTGAGAGAGATGATGACCGCTTCGCGAGGCGGTGGGCGGCGGCCAGCCCAAGCGCGCTG 1080  
DB 1027 CTGAGAGAGATGATGACCGCTTCGCGAGGCGGTGGGCGGCGGCCAGCCCAAGCGCGCTG 1086  
QY 1081 CTGCGCGAGGCGATGAGCGCAGCGCCCAACACACCGGTGATGTCAGAGAGCACTTCAAG 1140  
DB 1087 CTGCGCGAGGCGATGAGCGCAGCGCCCAACACACCGGTGATGTCAGAGAGCACTTCAAG 1146  
QY 1141 GCGCCCGGCGCATCTGTAAGTCTTCAACTGCGCGCAAGAGGCGCCACATCGCGCGCAAC 1200  
DB 1147 GCGCCCGGCGCATCTGTAAGTCTTCAACTGCGCGCAAGAGGCGCCACATCGCGCGCAAC 1206



QY 1201 TCCCGCCCCCCCCGCAAGAGGCTCTGTGAATGCGGCAAGAGGGCCACCAATGAAG 1260  
DB 1207 TCCCGCCCCCCCCGCAAGAGGCTCTGTGAATGCGGCAAGAGGGCCACCAATGAAG 1266  
QY 1261 GACTGCACCGAGCGCCCAAGGCTTCTCTGGGCAAGATCTGGCCCAAGCCCAAGGGCCGC 1320  
DB 1267 GACTGCACCGAGCGCCCAAGGCTTCTCTGGGCAAGATCTGGCCCAAGCCCAAGGGCCGC 1326  
QY 1321 CCGCGCACTTCTCTGCAAGAGCGCCCGAGCCCAAGCCCGCCCGGAGCTTCCGC 1380  
DB 1327 CCGCGCACTTCTCTGCAAGAGCGCCCGAGCCCAAGCCCGCCCGGAGCTTCCGC 1386  
QY 1381 TTGAGAGAGCAACCCCGCGGCAAGAGCAAGAGAGCCGAGACCTTGAACAGC 1440  
DB 1387 TTGAGAGAGCAACCCCGCGGCAAGAGCAAGAGAGAGCCGAGACCTTGAACAGC 1446  
QY 1441 CTGAAGAGCTTGTTCGCAAGCAACCCCTGAGCCAGTAA 1479  
DB 1447 CTGAAGAGCTTGTTCGCAAGCAACCCCTGAGCCAGTAA 1485

RESULT 6  
ACA03523

ID ACA03523 standard; DNA; 4419 BP.

AC ACA03523;

DT 22-MAY-2003 (first entry)

DE Synthetic DNA encoding immunogenic HIV peptide #6.

XX Immunogenic HIV polypeptide; human immunodeficiency virus; HIV; vaccine;

KW gene therapy; packaging cell line; humoral immune response;

KW cellular immune response; gene delivery vector; DNA immunisation; ds.

OS Synthetic.

XX WO2003004657-A1.

PD 16-JAN-2003.

PF 05-JUL-2002; 2002MO-US021421.

PR 05-JUL-2001; 2001US-0303192P.

PR 31-AUG-2001; 2001US-0316860P.

PR 16-JAN-2002; 2002US-0349728P.

PR 16-JAN-2002; 2002US-0349793P.

PR 16-JAN-2002; 2002US-0349871P.

PA (CHIR ) CHIRON CORP.

PI Zur Megede J, Barnett SW, Lian Y;

DR WPI; 2003-221602/21.

XX New synthetic polynucleotides encoding antigenic HIV type B and/or type C

PT polypeptides, useful as immunogenic compositions or vaccines for

PT generating humoral or cellular immune responses against HIV in a subject,

PT especially humans.

PS Example 1; Fig 11; 262pp; English.

XX The invention describes a synthetic polynucleotide encoding 2 or more

CC immunogenic HIV polypeptides, where at least 2 of the polypeptides are

CC derived from different HIV subtypes. The polynucleotide is useful for

CC immunisation, generation of packaging cell lines, or production of HIV

CC polypeptides. The polynucleotide and its encoded proteins are useful as

CC immunogenic compositions or vaccines for generating humoral or cellular

CC immune responses against HIV in a subject, or for inducing neutralising

CC antibodies against HIV. The gene delivery vector comprising the

CC polynucleotide is also useful for DNA immunisation of, or for generating

CC an immune response (e.g. a humoral or cellular immune response) in, a

CC subject such as a mammal, particularly a human. This sequence encodes a

CC human immunodeficiency virus immunogenic peptide  
XX Sequence 4419 BP; 980 A; 1582 C; 1371 G; 486 T; 0 U; 0 Other;

Query Match 99.9%; Score 1477.4; DB 7; Length 4419;  
Best Local Similarity 99.9%; Pred. No. 3.1e-185;  
Matches 1478; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGCCCGCCCAAGATCTCTGCGCGCGGCAAGCTGAGCGCTGGAGCGCATCCGC 60  
DB 7 ATGGGCGCCCGCCCAAGATCTCTGCGCGCGGCAAGCTGAGCGCTGGAGCGCATCCGC 66  
QY 61 CTGCGCGCCCGCGGCAAGAGTGTCTACATGATGAAGCACTGGTGTGGGCCAGCCGCGAG 120  
DB 67 CTGCGCGCCCGCGGCAAGAGTGTCTACATGATGAAGCACTGGTGTGGGCCAGCCGCGAG 126  
QY 121 CTGAGAAAGTTCCCTCTGAACCCCGGCTGTGAGAGACCAAGGAGGCTGCAAGCAGATC 180  
DB 127 CTGAGAAAGTTCCCTCTGAACCCCGGCTGTGAGAGACCAAGGAGGCTGCAAGCAGATC 186  
QY 181 ATCCGCCAGCTGCAACCCCGGCTGCAAGCCGCGGCAAGAGAGCTGAAGAGCTGTTCAAC 240  
DB 187 ATCCGCCAGCTGCAACCCCGGCTGCAAGCCGCGGCAAGAGAGCTGAAGAGCTGTTCAAC 246  
QY 241 ACCGTGGCCACCTGTACTGCGGTGACAGAGAGATGAGAGTCCGCGCAACCAAGAGGCC 300  
DB 247 ACCGTGGCCACCTGTACTGCGGTGACAGAGAGATGAGAGTCCGCGCAACCAAGAGGCC 306  
QY 301 CTGACAAAGATCGAGAGAGAGCAAACTGTCAGCAAGAGATCCAGAGCGCGAGGCC 360  
DB 307 CTGACAAAGATCGAGAGAGAGCAAACTGTCAGCAAGAGATCCAGAGCGCGAGGCC 366  
QY 361 GCCGCAAGGGCAAGGTGAGCGCAAACTGTCAGCAAGAGATCCAGAGCGCGAGGCCAGATG 420  
DB 367 GCCGCAAGGGCAAGGTGAGCGCAAACTGTCAGCAAGAGATCCAGAGCGCGAGGCCAGATG 426  
QY 421 GTGACACGAGCCATCAGCCCGGCAACCTGTAACGCTGGGTGAAGGTGATCGAGAGAAAG 480  
DB 427 GTGACACGAGCCATCAGCCCGGCAACCTGTAACGCTGGGTGAAGGTGATCGAGAGAAAG 486  
QY 481 GCCTTCAAGCCCGAGGTGATCCCATGTTCAACGCTGGGTGAAGGTGATCGAGAGAAAG 540  
DB 487 GCCTTCAAGCCCGAGGTGATCCCATGTTCAACGCTGGGTGAAGGTGATCGAGAGAAAG 546  
QY 541 GACCTGAACAGATGTTGAACACCGTGGCGGCGCACAGGCGGCGCATGCAATGCTGAAG 600  
DB 547 GACCTGAACAGATGTTGAACACCGTGGCGGCGCACAGGCGGCGCATGCAATGCTGAAG 606  
QY 601 GACACCATCAACGAGAGCGCGGAGTGGACCGGCGTGCACCCCGTGCACCGCGGCGGCGG 660  
DB 607 GACACCATCAACGAGAGCGCGGAGTGGACCGGCGTGCACCCCGTGCACCGCGGCGGCGG 666  
QY 661 ATGCCCCCGGCGAGATGCGCGAGCGCCCGGCGAGCGACATGCGCGGCGACCAAGCACCC 720  
DB 667 ATGCCCCCGGCGAGATGCGCGAGCGCCCGGCGAGCGACATGCGCGGCGACCAAGCACCC 726  
QY 721 CTGAGAGAGCAGATGCGCTGATGACCAAGCAACCCCGCATCCCGTGGGCGCATCTAC 780  
DB 727 CTGAGAGAGCAGATGCGCTGATGACCAAGCAACCCCGCATCCCGTGGGCGCATCTAC 786  
QY 781 AAGCGGTGATCATCTGGGCTGAACAAGATCGTGGAGATGAAGCCCGGTGAGCATC 840  
DB 787 AAGCGGTGATCATCTGGGCTGAACAAGATCGTGGAGATGAAGCCCGGTGAGCATC 846  
QY 841 CTGAGACATCAAGCAGAGGCGGCAAGAGCGCTTCCGCGACTACGTTCAAG 900  
DB 847 CTGAGACATCAAGCAGAGGCGGCAAGAGCGCTTCCGCGACTACGTTCAAG 906  
QY 901 ACCCTGCGCGCGAGCAGAGCAACCGAGAGTGAAGACTGATGAACGACACCTGCTG 960  
DB 907 ACCCTGCGCGCGAGCAGAGCAACCGAGAGTGAAGACTGATGAACGACACCTGCTG 966  
QY 961 GTGCAAGAGCGCAACCCGACTGCAAGACCATCTGCGCGCTTCCGCGCGCGGCGGCGAGC 1020



```
|||||
DB 967 GTGAGAAAGCGCAACCCGACTGCAAGACCATCTGCGCGCTCTGCGCCCGCGCCAGC 1026
OY 1021 CTGAGAGAGATGATGACCGCTGCGAGAGCGGTGGCGGCCCCAGCCACAAGCCCGCTG 1080
|||||
DB 1027 CTGAGAGAGATGATGACCGCTGCGAGAGCGGTGGCGGCCCCAGCCACAAGCCCGCTG 1086
OY 1081 CTGCGCGAGCGGATGAGCGGCGCAACACAGCGGTGATGATGAGAGAGAGCAACTTCAAG 1140
|||||
DB 1087 CTGCGCGAGCGGATGAGCGGCGCAACACAGCGGTGATGATGAGAGAGAGCAACTTCAAG 1146
OY 1141 GCGCGCGCGGCGGATGCTCAAGTGTCTCAACTGCGGCAAGAGGCGCAATCGCCGCAAC 1200
|||||
DB 1147 GCGCGCGCGGCGGATGCTCAAGTGTCTCAACTGCGGCAAGAGGCGCAATCGCCGCAAC 1206
OY 1201 TCGCGCGCGCGCGCGCAAGAGGCGTGTGAAAGTGGCGCAAGAGGCGCAACAGATGAAG 1260
|||||
DB 1207 TCGCGCGCGCGCGCGCAAGAGGCGTGTGAAAGTGGCGCAAGAGGCGCAACAGATGAAG 1266
OY 1261 GACTGCAACCGAGCGCCAGGCAACTTCTGTGGGCAAGATCTGCGGCAACCAAGGCGCGC 1320
|||||
DB 1267 GACTGCAACCGAGCGCCAGGCAACTTCTGTGGGCAAGATCTGCGGCAACCAAGGCGCGC 1326
OY 1321 CCGCGCAACTTCTGTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 1380
|||||
DB 1327 CCGCGCAACTTCTGTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 1386
OY 1381 TTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
|||||
DB 1387 TTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1446
OY 1441 CTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1479
|||||
DB 1447 CTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1485
```

```
RESULT 7
ADCI3240
ID ADCI3240 standard; DNA; 4419 BP.
AC ADCI3240;
DT 18-DEC-2003 (first entry)
DE DNA of HIV construct GagRTmutatRevNef_C SEQ ID NO 19.
KM expression cassette; HIV Gag; Env; Int; Nef; p15RnaseH; Pol; Tat; Prot;
XX Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; ds.
OS Human immunodeficiency virus.
PN WO2003004620-A2.
PD 16-JAN-2003.
PF 05-JUL-2002; 2002WO-US021420.
PR 05-JUL-2001; 2001US-0303192P.
PR 31-AUG-2001; 2001US-0316860P.
PR 16-JAN-2002; 2002US-0349871P.
PA (CHIR ) CHIRON CORP.
PA (UYST-) UNIV STRLEMBOSCH.
PI Zur Megele J, Barnett SW, Lian Y, Engelbrecht S, Van Rensburg BJ;
DR WPI; 2003-221593/21.
XX New expression cassette comprising a polynucleotide sequence encoding a
PT polypeptide including an HIV Gag, Env, Int, Nef, p15RnaseH, Pol, Tat,
PT Prot, or Rev polypeptide, useful for immunization, or generating
XX packaging cell lines.
```

```
PS Disclosure; Fig 16; 301pp; English.
XX
CC The invention relates to a novel expression cassette comprising a
CC polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,
CC Int, Nef, p15RnaseH, Pol, Tat, Prot, or Rev polypeptide. The novel
CC expression cassette can be used to treat HIV type C by gene therapy or
CC used in the development of a vaccine. The gene delivery vector is
CC administered intramuscularly, intramuscularly, intranasally,
CC subcutaneously, intradermally, transdermally, intravenously,
CC intrarectally, orally or intravenously. The expression cassette is useful
CC for immunisation, generating packaging cell lines and producing HIV
CC polypeptides. This polynucleotide sequence represents the DNA of an HIV
CC type C related sequence of the invention.
XX
SQ Sequence 4419 BP; 980 A; 1582 C; 1371 G; 486 T; 0 U; 0 Other;

Query Match 99.9%; Score 1477.4; DB 9; Length 4419;
Best Local Similarity 99.9%; Pred. No. 3.1e-185;
Matches 1478; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ATGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
DB 7 ATGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 66
OY 61 CTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
DB 67 CTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 126
OY 121 CTGAGAGAGTTCGCGCTTGAACCGCGCGCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAG 180
DB 127 CTGAGAGAGTTCGCGCTTGAACCGCGCGCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAG 186
OY 181 ATCGCGCGAGCTGACACCGCGCGCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
DB 187 ATCGCGCGAGCTGACACCGCGCGCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 246
OY 241 ACCGTGGCGCAACCTCTGTACTGCGGTGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
DB 247 ACCGTGGCGCAACCTCTGTACTGCGGTGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 306
OY 301 CTGAGCAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
DB 307 CTGAGCAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 366
OY 361 GCGGCAAGAGGCGAAGGTGAGCGCGAAGTACCCTGTCAGAGAGAGAGAGAGAGAGAGAGAG 420
DB 367 GCGGCAAGAGGCGAAGGTGAGCGCGAAGTACCCTGTCAGAGAGAGAGAGAGAGAGAGAGAG 426
OY 421 GTGCAACAGGCGCATGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
DB 427 GTGCAACAGGCGCATGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 486
OY 481 GCCTTCAGCGCGCGAGGTGATCCCATGTTCAACCGCGCTGAGCGAGGCGCGCAACCGCGCGAG 540
DB 487 GCCTTCAGCGCGCGAGGTGATCCCATGTTCAACCGCGCTGAGCGAGGCGCGCAACCGCGCGAG 546
OY 541 GACTGAACACGATGTTGAACACCGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
DB 547 GACTGAACACGATGTTGAACACCGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 606
OY 601 GACACCATCAACGAGAGAGCGCGCGAGTGGAGACCGCGGTGCAACCGCGGTGCAACCGCGCGCG 660
DB 607 GACACCATCAACGAGAGAGCGCGCGAGTGGAGACCGCGGTGCAACCGCGGTGCAACCGCGCGCG 666
OY 661 ATGCGCGCGCGCGCGAGATGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
DB 667 ATGCGCGCGCGCGCGAGATGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 726
OY 721 CTGAGAGAGAGATGCGCTGTGATGACGAGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 780
DB 727 CTGAGAGAGAGATGCGCTGTGATGACGAGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 786
OY 781 AAGCGGTGATCATCTGCGCGCTGAACCAAGATCTGCGGATGTACAGCGCGCGGTGAGCATC 840
```

```
DB 787 AAGCGTGATCATCTGGGCTGAACAAGATGTGCGATGTACAGCCCGTGAGCATC 846
QY 841 CTGACATCAAGCAGGGCCCCAAGAGGCCCTTCGGGACTACGTGAGCCGCTTCAAG 900
DB 847 CTGACATCAAGCAGGGCCCCAAGAGGCCCTTCGGGACTACGTGAGCCGCTTCAAG 906
QY 901 ACCCTGCGCGGAGCAGAGCAACCCAGAGGTGAAGACTGATGACCGACACCTGCTG 960
DB 907 ACCCTGCGCGGAGCAGAGCAACCCAGAGGTGAAGACTGATGACCGACACCTGCTG 966
QY 961 GTGCAAGACCGCAACCCGACTGCAAGACCATCTGCGGCTCTCGGCCCGGCGCAGC 1020
DB 967 GTGCAAGACCGCAACCCGACTGCAAGACCATCTGCGGCTCTCGGCCCGGCGCAGC 1026
QY 1021 CTGAGAGATGATGACCGCTGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
DB 1027 CTGAGAGATGATGACCGCTGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1086
QY 1081 CTGCGCGAGCGATGAGCCGACCAACCAACGCTGATGATGACGAAGCACTTCAAG 1140
DB 1087 CTGCGCGAGCGATGAGCCGACCAACCAACGCTGATGATGACGAAGCACTTCAAG 1146
QY 1141 GGGCGCGCGGCGCATGCTCAAGTGTCTCAACTGCGGCGAAGAGGCGGCACATCGCCGCAAC 1200
DB 1147 GGGCGCGCGGCGCATGCTCAAGTGTCTCAACTGCGGCGAAGAGGCGGCACATCGCCGCAAC 1206
QY 1201 TGCGCGCGCGCGCGCAAGAGGCGCTGCGAAGTGTGCGGCAAGAGGCGGCACATGAAAG 1260
DB 1207 TGCGCGCGCGCGCGCAAGAGGCGCTGCGAAGTGTGCGGCAAGAGGCGGCACATGAAAG 1266
QY 1261 GACTGACCGAGCGCGCAAGGCGCACTTCTGCGGCAAGATCTGCGCCAGCCACAAAGGCGCGC 1320
DB 1267 GACTGACCGAGCGCGCAAGGCGCACTTCTGCGGCAAGATCTGCGCCAGCCACAAAGGCGCGC 1326
QY 1321 CTGCGCACTTCTGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTCGC 1380
DB 1327 CTGCGCACTTCTGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTCGC 1386
QY 1381 TTGAGAGAGACCAACCCCGCGCGCAAGAGCAGAGCAAGACCGGAGACCTTGAACAGC 1440
DB 1387 TTGAGAGAGACCAACCCCGCGCGCAAGAGCAGAGCAAGACCGGAGACCTTGAACAGC 1446
QY 1441 CTGAAGAGCTGTTGGGCAAGCAACCCCTGAGCCAGTAA 1479
DB 1447 CTGAAGAGCTGTTGGGCAAGCAACCCCTGAGCCAGTAA 1485
```

RESULT 8  
ADCl3257  
ID ADCl3257 standard; DNA; 4615 BP.

```
XX AC ADCl3257;  
XX DT 18-DEC-2003 (first entry)  
DB DNA of HIV construct gp160mod-TV1-dV2-gagmod-BW965 SEQ ID NO 36.  
XX expression cassette; HIV Gag; Env; Int; Nef; p15RnaseH; Pol; Tat; Proct;  
KM Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; ds.  
XX Human immunodeficiency virus.  
XX OS  
XX PN WO2003004620-A2.  
XX PD 16-JAN-2003.  
XX PP 05-JUL-2002; 2002WO-US021420.  
XX PR 05-JUL-2001; 2001US-0303192P.  
XX PR 31-AUG-2001; 2001US-0316860P.  
XX PR 16-JAN-2002; 2002US-0349871P.  
XX
```

```
PA (CHIR ) CHIRON CORP.  
PA (UYST-) UNIV STRLEBOSCH.  
XX Zur Megede J, Barnett SW, Lian Y, Engelbrecht S, Van Rensburg EJ;  
XX WPI; 2003-221593/21.  
XX  
PT New expression cassette comprising a polynucleotide sequence encoding a  
PT polypeptide including an HIV Gag, Env, Int, Nef, p15RnaseH, Pol, Tat,  
PT Prot, or Rev polypeptide, useful for immunization, or generating  
PT packaging cell lines.  
XX  
PS Disclosure; Fig 33; 301pp; English.  
XX  
CC The invention relates to a novel expression cassette comprising a  
CC polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,  
CC Int, Nef, p15RnaseH, Pol, Tat, Prot, or Rev polypeptide. The novel  
CC expression cassette can be used to treat HIV type C by gene therapy or  
CC used in the development of a vaccine. The gene delivery vector is  
CC administered intramuscularly, intravenously, intranasally,  
CC subcutaneously, intradermally, transdermally, intravaginally,  
CC intrarectally, orally or intravenously. The expression cassette is useful  
CC for immunisation, generating packaging cell lines and producing HIV  
CC polypeptides. This polynucleotide sequence represents the DNA of an HIV  
CC Type C related sequence of the invention.  
XX  
SQ Sequence 4615 BP; 1030 A; 1587 C; 1354 G; 644 T; 0 U; 0 Other;  
  
Query Match 99.9%; Score 1477.4; DB 9; Length 4615;  
Best Local Similarity 99.9%; Pred. No. 3.1e-185;  
Matches 1478; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 ATGGGCGCGCGCGCGCGCATCTGCGGCGCGCGCAAGCTGAGCGCTGGAGCGCATCCGC 60  
DB 3137 ATGGGCGCGCGCGCGCGCGCATCTGCGGCGCGCGCAAGCTGAGCGCTGGAGCGCATCCGC 3196  
QY 61 CTGCGCGCGCGCGCGCAAGAGTGTCTACATGATGAAGCACCTGTGTGCGCGCGCGAG 120  
DB 3197 CTGCGCGCGCGCGCGCAAGAGTGTCTACATGATGAAGCACCTGTGTGCGCGCGCGAG 3256  
QY 121 CTGAGAGATTCGCTGAAACCCCGCGCTGCTGAGAGACCGAGGCGCTGCAAGCATC 180  
DB 3257 CTGAGAGATTCGCTGAAACCCCGCGCTGCTGAGAGACCGAGGCGCTGCAAGCATC 3316  
QY 181 ATCCGCGAGCTGACCCCGCGCTGACAGCCGCGCAGGAGCTGAAGAGCTGTTCAAC 240  
DB 3317 ATCCGCGAGCTGACCCCGCGCTGACAGCCGCGCAGGAGCTGAAGAGCTGTTCAAC 3376  
QY 241 ACCGTGCGCACCTGTACTGCTGTCAGAGAGATGAGGTCCGACACCAAGAGGCC 300  
DB 3377 ACCGTGCGCACCTGTACTGCTGTCAGAGAGATGAGGTCCGACACCAAGAGGCC 3436  
QY 301 CTGACAAAGATCGAGAGAGCAAGAACAGTCCAGAGAGATCCAGCAGCGCGAGGCC 360  
DB 3437 CTGACAAAGATCGAGAGAGCAAGAACAGTCCAGAGAGATCCAGCAGCGCGAGGCC 3496  
QY 361 GCGGACAAAGGCGCAAGGTGAGCCAGAACTACCCCATGCTGTCAGAACTGCAAGGCCAGATG 420  
DB 3497 GCGGACAAAGGCGCAAGGTGAGCCAGAACTACCCCATGCTGTCAGAACTGCAAGGCCAGATG 3556  
QY 421 GTGACACAGGCGCATCAGCCCGCGCACCTGTAAGCGCTGGTGAAGGTGATCGAGAGAAG 480  
DB 3557 GTGACACAGGCGCATCAGCCCGCGCACCTGTAAGCGCTGGTGAAGGTGATCGAGAGAAG 3616  
QY 481 GCCTTCAAGCGCGAGGTGATCCCATATGTTCAACCGCCCTGAGGAGGCGCGCCAGCCCCCAG 540  
DB 3617 GCCTTCAAGCGCGAGGTGATCCCATATGTTCAACCGCCCTGAGGAGGCGCGCCAGCCCCCAG 3676  
QY 541 GACCTGAACAGATGTTGAACACCGTGGGCGCCACAGCGCGCATGACATGCTGAAG 600  
DB 3677 GACCTGAACAGATGTTGAACACCGTGGGCGCCACAGCGCGCATGACATGCTGAAG 3736  
QY 601 GACACCATCAAGAGAGCGCGCGAGTGGAGACCGCGTGAACCCCGTGAACCGCGCGCCCC 660
```



```
Db 3644 GTGACACGAGCCATCAGCCCCCGACCCCTGAAAGCTGGTGAAGTGAATGAGAGAAAG 3703
Qy 481 GCGTTGAGCCCCGAGGTGATCCCATGTTCAACGCGCTGAAGAGGCGCCACCCCGCAG 540
Db 3704 GCCTTCAGCCCCGAGGTGATCCCATGTTCAACGCGCTGAAGAGGCGCCACCCCGCAG 3763
Qy 541 GACCTGAACAGATGTTGAACAGGTGGGCGCCACAGGCGCCATGAGATGCTGAAG 600
Db 3764 GACCTGAACAGATGTTGAACAGGTGGGCGCCACAGGCGCCATGAGATGCTGAAG 3823
Qy 601 GACACCATCAACGAGAGGCGCGGAGTGGGACCGCTGCAACCCCGTGCAGCGCGCGCC 660
Db 3824 GACACCATCAACGAGAGGCGCGGAGTGGGACCGCTGCAACCCCGTGCAGCGCGCGCC 3883
Qy 661 ATGCCCCCGGCGAGATGCGCGAGCGCCCGCGCGAGCAATCGCGCGCACCAACGACCC 720
Db 3884 ATGCCCCCGGCGAGATGCGCGAGCGCCCGCGCGAGCAATCGCGCGCACCAACGACCC 3943
Qy 721 CTGCAAGACAGATCGCTGGAATGACCAAGAACCCCGCATCCCGTGGCGCATCTTAC 780
Db 3944 CTGCAAGACAGATCGCTGGAATGACCAAGAACCCCGCATCCCGTGGCGCATCTTAC 4003
Qy 781 AAGCGGTGATCATCTGCGCTGAACAAGATCGTGGGATGTACAGCCCGGTGAGCATC 840
Db 4004 AAGCGGTGATCATCTGCGCTGAACAAGATCGTGGGATGTACAGCCCGGTGAGCATC 4063
Qy 841 CTGACATCAAGCAGGCGCCCAAGAGCGCTTCCGCGACTACGTGAGCGCTTCTTCAAG 900
Db 4064 CTGACATCAAGCAGGCGCCCAAGAGCGCTTCCGCGACTACGTGAGCGCTTCTTCAAG 4123
Qy 901 ACCCTGCGCGCGGAGCAGACACCAAGAGGTGAAGAACTGATGACCGCACCTCTGCTG 960
Db 4124 ACCCTGCGCGCGGAGCAGACACCAAGAGGTGAAGAACTGATGACCGCACCTCTGCTG 4183
Qy 961 GTGCAAGACGCCCAACCCCGACTGCAAGACCATCTGCGCGCTCTGCGCGCGCGCGCAGC 1020
Db 4184 GTGCAAGACGCCCAACCCCGACTGCAAGACCATCTGCGCGCTCTGCGCGCGCGCGCAGC 4243
Qy 1021 CTGAGAGAGATGATGACCGCTTGCAGAGGCGTGGCGCGCCCGCACAGCCAGCGCGCGT 1080
Db 4244 CTGAGAGAGATGATGACCGCTTGCAGAGGCGTGGCGCGCCCGCACAGCGCGCGCGT 4303
Qy 1081 CTGCGCGAGCGGATGAGCGGCGCCCAACCAAGCGTGTATGATGCAAGAGACAACTTCAAG 1140
Db 4304 CTGCGCGAGCGGATGAGCGGCGCCCAACCAAGCGTGTATGATGCAAGAGACAACTTCAAG 4363
Qy 1141 GCGCGCGCGCGCATCTGTAAGTGTCTCAACTGCGCGCAAGAGGCGCCCATGCGCGCGCAAC 1200
Db 4364 GCGCGCGCGCGCATCTGTAAGTGTCTCAACTGCGCGCAAGAGGCGCCCATGCGCGCGCAAC 4423
Qy 1201 TGCGCGCGCGCGCGCAAGAGGCGCTGTGAAAGTGCAGGCAAGAGGCGCCACAGATGAAG 1260
Db 4424 TGCGCGCGCGCGCGCAAGAGGCGCTGTGAAAGTGCAGGCAAGAGGCGCCACAGATGAAG 4483
Qy 1261 GACTGCACCGAGCGCCAGGCGCAACTTCTGCGCGCAAGATCTGCGCCAGCCACAGAGGCGCGC 1320
Db 4484 GACTGCACCGAGCGCCAGGCGCAACTTCTGCGCGCAAGATCTGCGCCAGCCACAGAGGCGCGC 4543
Qy 1321 CCGCGCAACTTCTGCAAGCGCGCGCGCGAGCGCCACCGCGCGCGCGCGAGAGCTTCCGC 1380
Db 4544 CCGCGCAACTTCTGCAAGCGCGCGCGCGAGCGCCACCGCGCGCGCGCGAGAGCTTCCGC 4603
Qy 1381 TTGAGAGAGACCAACCCCGCGCGCAAGAGCAGAGAGCAAGGACCGCGAGACCTTGACAGC 1440
Db 4604 TTGAGAGAGACCAACCCCGCGCGCAAGAGCAGAGAGCAAGGACCGCGAGACCTTGACAGC 4663
Qy 1441 CTGAAGAGCTGTTCGGCAAGACCCCGCTGAGCCAGTAA 1479
Db 4664 CTGAAGAGCTGTTCGGCAAGACCCCGCTGAGCCAGTAA 4702
```

RESULT 10

```
ACA03522
ID ACA03522 standard; DNA; 4716 BP.
XX
AC ACA03522;
XX
DT 22-MAY-2003 (first entry)
XX
DE Synthetic DNA encoding immunogenic HIV peptide #5.
XX
KW Immunogenic HIV polypeptide; human immunodeficiency virus; HIV; vaccine;
KW gene therapy; packaging cell line; humoral immune response;
KW cellular immune response; gene delivery vector; DNA immunisation; ds.
XX
OS Synthetic.
XX
PN WO2003004657-A1.
XX
PD 16-JAN-2003.
XX
PF 05-JUL-2002; 2002MO-US021421.
XX
PR 05-JUL-2001; 2001US-0303192P.
PR 31-AUG-2001; 2001US-0316860P.
PR 16-JAN-2002; 2002US-0349728P.
PR 16-JAN-2002; 2002US-0349793P.
PR 16-JAN-2002; 2002US-0349871P.
XX
PA (CHIR ) CHIRON CORP.
XX
PI Zur Megele J, Barnett SW, Lian Y;
XX
DR WPI; 2003-221602/21.
XX
PT New synthetic polynucleotides encoding antigenic HIV type B and/or type C
PT polypeptides, useful as immunogenic compositions or vaccines for
PT generating humoral or cellular immune responses against HIV in a subject,
PT especially humans.
XX
PS Example 1; Fig 10; 262pp; English.
XX
CC The invention describes a synthetic polynucleotide encoding 2 or more
CC immunogenic HIV polypeptides, where at least 2 of the polypeptides are
CC derived from different HIV subtypes. The polynucleotide is useful for
CC immunisation, generation of packaging cell lines, or production of HIV
CC polypeptides. The polynucleotide and its encoded proteins are useful as
CC immunogenic compositions or vaccines for generating humoral or cellular
CC immune responses against HIV in a subject, or for inducing neutralising
CC antibodies against HIV. The gene delivery vector comprising the
CC polynucleotide is also useful for DNA immunisation of, or for generating
CC an immune response (e.g. a humoral or cellular immune response) in, a
CC subject such as a mammal, particularly a human. This sequence encodes a
CC human immunodeficiency virus immunogenic peptide
XX
SQ Sequence 4716 BP; 1041 A; 1684 C; 1463 G; 528 T; 0 U; 0 Other;
XX
Query Match 99.9%; Score 1477.4; DB 7; Length 4716;
Best Local Similarity 99.9%; Pred. No. 3,1e-185;
Matches 1478; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ATGGGCGCGCGCGCGCAAGCATCTGCGCGGCGCAAGCTGAGCGCTTGGAGCGCATCCGC 60
Db 7 ATGGGCGCGCGCGCGCAAGCATCTGCGCGGCGCAAGCTGAGCGCTTGGAGCGCATCCGC 66
Qy 61 CTGCGCGCGCGCGCGCAAGAGTGTCTACATGATGAAGCACTGTGTGTGGCCAGCGCGAG 120
Db 67 CTGCGCGCGCGCGCGCAAGAGTGTCTACATGATGAAGCACTGTGTGTGGCCAGCGCGAG 126
Qy 121 CTGAGAGAGTTGCGCTGAACCCCGGCTGTGAGAGCCAGCGAGGGCTGCAAGCAGATC 180
Db 127 CTGAGAGAGTTGCGCTGAACCCCGGCTGTGAGAGCCAGCGAGGGCTGCAAGCAGATC 186
Qy 181 ATCGCGCAGCTGCACCCCGCGCTGCAAGCGCGCAGCGAGAGCTGAAGAGCTGTTCAC 240
Db ATCGCGCAGCTGCACCCCGCGCTGCAAGCGCGCAGCGAGAGCTGAAGAGCTGTTCAC 240
```



```

Db      187 ATCCGCCAGCTGCAACCCCGCCCTGCAAGACCCGCGACGAGAGCTGAAGACCTGTTCAAC 246
Qy      241 ACCGTGGCCACCTCTGTACTGCTGTGCAACGAGAAGATGAGGTCCCGCACACCAAGAGGCC 300
Db      247 ACCGTGGCCACCTCTGTACTGCTGTGCAACGAGAAGATGAGGTCCCGCACACCAAGAGGCC 306
Qy      301 CTGGAACAAGATGAGAGAGAGAGCAACAAGTCCAGCAAGAATCCAGAGCCGAGGCC 360
Db      307 CTGGAACAAGATGAGAGAGAGAGCAACAAGTCCAGCAAGAATCCAGAGCCGAGGCC 366
Qy      361 GCCGACAAGGGGCAAGTGAAGCCCAACTACCCCATCGTGCAGAACTGCAAGGCCAGATG 420
Db      367 GCCGACAAGGGGCAAGTGAAGCCCAACTACCCCATCGTGCAGAACTGCAAGGCCAGATG 426
Qy      421 GTGCACCAAGGCAATCAGCCCGCGCAACCTGAACGCTTGGTGAAGGTGATCGAGAGAG 480
Db      427 GTGCACCAAGGCAATCAGCCCGCGCAACCTGAACGCTTGGTGAAGGTGATCGAGAGAG 486
Qy      481 GCCTTCAGCCCGGAGGTGATCCCATGTTCAACGCTTGAAGCGGCGCCACCCCCAG 540
Db      487 GCCTTCAGCCCGGAGGTGATCCCATGTTCAACGCTTGAAGCGGCGCCACCCCCAG 546
Qy      541 GACCTGACACGATGTTGAACAACCTGTGGCGGCGCAACAGGCGCCATGCAAGTGTGAAG 600
Db      547 GACCTGACACGATGTTGAACAACCTGTGGCGGCGCAACAGGCGCCATGCAAGTGTGAAG 606
Qy      601 GACACCATCAACGAGAGAGGCGCGGAGTGGAGCTGCAACCCCGTGCACCGCGGCCCC 660
Db      607 GACACCATCAACGAGAGAGGCGCGGAGTGGAGCTGCAACCCCGTGCACCGCGGCCCC 666
Qy      661 ATCGCCCCCGGCGAGATGCGCGAGCCCGCGCGGCGAGCAATCGCGCGGCAACCAAGCAC 720
Db      667 ATCGCCCCCGGCGAGATGCGCGAGCCCGCGCGGCGAGCAATCGCGCGGCAACCAAGCAC 726
Qy      721 CTGCAAGAGCAGATCGCTGTGATGACCAAGCAACCCCGCATCCCGTGGCGCAATCTAC 780
Db      727 CTGCAAGAGCAGATCGCTGTGATGACCAAGCAACCCCGCATCCCGTGGCGCAATCTAC 786
Qy      781 AAGCGGTGATCATCTGGGCTGAAACAAGATGTCGGGATGTAACAGCCCGTGAACATC 840
Db      787 AAGCGGTGATCATCTGGGCTGAAACAAGATGTCGGGATGTAACAGCCCGTGAACATC 846
Qy      841 CTGACATCAAGCAAGGCGCCCAAGAGAGCCCTTCGCGACTACGTGGAACCGCTTCTCAAG 900
Db      847 CTGACATCAAGCAAGGCGCCCAAGAGAGCCCTTCGCGACTACGTGGAACCGCTTCTCAAG 906
Qy      901 ACCCTGCGCGCGGAGCAAGAGCAACCAAGAGGTGAAGAATGTGATGACCGACCTGCTG 960
Db      907 ACCCTGCGCGCGGAGCAAGAGCAACCAAGAGGTGAAGAATGTGATGACCGACCTGCTG 966
Qy      961 GTGCAGAACGCCAACCCCGACTGCAAGACCATCTGCGCGCTCTCGGCCCGGCGCCAGC 1020
Db      967 GTGCAGAACGCCAACCCCGACTGCAAGACCATCTGCGCGCTCTCGGCCCGGCGCCAGC 1026
Qy      1021 CTGAGAGAGATGATGACCGCTGCGAGGCGGTGGGCGGCGCCCAAGCCCAAGGCGCGGTG 1080
Db      1027 CTGAGAGAGATGATGACCGCTGCGAGGCGGTGGGCGGCGCCCAAGCCCAAGGCGCGGTG 1086
Qy      1081 CTGGCCGAGCGATGAGCCAGGCGCAACACCAAGCGTGTATGATGCAAGAGCAACTTCAMG 1140
Db      1087 CTGGCCGAGCGATGAGCCAGGCGCAACACCAAGCGTGTATGATGCAAGAGCAACTTCAMG 1146
Qy      1141 GGGCCCCCGGCGATGTCGAAGTCTTCAACTGCGGCAAGAGGCGCCACATGCGCGCAAC 1200
Db      1147 GGGCCCCCGGCGATGTCGAAGTCTTCAACTGCGGCAAGAGGCGCCACATGCGCGCAAC 1206
Qy      1201 TGCCGCGCGCGCGCAAGAGGGGCTGTGGAAGTGCAGCAAGAGGGGCGCCACCATGAAG 1260
Db      1207 TGCCGCGCGCGCGCAAGAGGGGCTGTGGAAGTGCAGCAAGAGGGGCGCCACCATGAAG 1266
Qy      1261 GACTGCACCGAGCGCGCAAGGCCAATTCTTGCGCAAGATCTGCGCCAGCCCAAGGGCGCG 1320
Db      1267 GACTGCACCGAGCGCGCAAGGCCAATTCTTGCGCAAGATCTGCGCCAGCCCAAGGGCGCG 1326

```

```

Qy      1321 CCGGCAACTTCTCTGAGAGCGCGCGCGAGCCCAACCGCCCCCGCGGAGAGCTCCGC 1380
Db      1327 CCGGCAACTTCTCTGAGAGCGCGCGCGAGCCCAACCGCCCCCGCGGAGAGCTCCGC 1386
Qy      1381 TTGAGGAGACCAACCCCGCGGCAAGAGCAAGAGCAAGAGACCGCGGAGACCTGACAGC 1440
Db      1387 TTGAGGAGACCAACCCCGCGGCAAGAGCAAGAGCAAGAGACCGCGGAGACCTGACAGC 1446
Qy      1441 CTGAAGAGCTGTTCGCAAGCAACCCCTGAGCCAGTAA 1479
Db      1447 CTGAAGAGCTGTTCGCAAGCAACCCCTGAGCCAGAAA 1485

```

## RESULT 11

ADCl3238  
ID ADCl3238 standard; DNA; 4716 BP.

XX  
AC ADCl3238;

DT 18-DEC-2003 (first entry)

DE DNA of HIV construct GagProIntartmutTatRevNef\_C SEQ ID NO 17.

XX  
KW expression cassette; HIV Gag; Env; Int; Nef; p15RnaseH; Pol; Tat; Prot;  
KW Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; ds.

XX  
OS Human immunodeficiency virus.

XX  
PN WO2003004620-A2.

XX  
PD 16-JAN-2003.

XX  
PF 05-JUL-2002; 2002WO-US021420.

XX  
PR 05-JUL-2001; 2001US-0303192P.

XX  
PR 31-AUG-2001; 2001US-0316860P.

XX  
PR 16-JAN-2002; 2002US-0349871P.

XX  
PA (CHIR ) CHIRON CORP.

XX  
PA (UYST-) UNIV STELBENBOSCH.

XX  
PI Zur Megede J, Barnett SW, Lian Y, Engelbrecht S, Van Rensburg EJ;

XX  
DR WPI; 2003-221593/21.

XX  
PT New expression cassette comprising a polynucleotide sequence encoding a  
PT polypeptide including an HIV Gag, Env, Int, Nef, p15RnaseH, Pol, Tat,  
PT Prot, or Rev polypeptide, useful for immunization, or generating  
PT packaging cell lines.

XX  
PS Disclosure; Fig 14; 301pp; English.

XX  
CC The invention relates to a novel expression cassette comprising a  
CC polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,  
CC Int, Nef, p15RnaseH, Pol, Tat, Prot, or Rev polypeptide. The novel  
CC expression cassette can be used to treat HIV type C by gene therapy or  
CC used in the development of a vaccine. The gene delivery vector is  
CC administered intramuscularly, intravenously, intranasally,  
CC subcutaneously, intradermally, transdermally, intravaginally,  
CC intrarectally, orally or intravenously. The expression cassette is useful  
CC for immunisation, generating packaging cell lines and producing HIV  
CC polypeptides. This polynucleotide sequence represents the DNA of an HIV  
CC Type C related sequence of the invention.

XX  
SQ Sequence 4716 BP; 1041 A; 1684 C; 1463 G; 528 T; 0 U; 0 Other;

Query Match 99.9%; Score 1477.4; DB 9; Length 4716;  
Best Local Similarity 99.9%; Pred. No. 3.1e-185;  
Matches 1478; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGGGCGCGCGCGCAGCATCTGCGCGCGCGCAAGCTGACGCTGAGCGCATCCGC 60  
|||||



```

Db      7 ATGGGCGCCCGCGCCAGCATCTCGCGGCGCAAGCTGAGCGCCTGGAGCGCATCCGC 66
Qy      61 CTGCGCCCGCGCGCAAGAGTGTACTATGATGAAGCACTTGTGTGGCCAGCCGCGAG 120
        |||
Db      67 CTGCGCCCGCGCGCAAGAGTGTACTATGATGAAGCACTTGTGTGGCCAGCCGCGAG 126
Qy      121 CTGAGAGAGTTGCGCCCTGAACCTCCGCGCTGTCTGAGAGCCAGCGAGGGGTGCAAGCAGATC 180
        |||
Db      127 CTGAGAGAGTTGCGCCCTGAACCTCCGCGCTGTCTGAGAGCCAGCGAGGGGTGCAAGCAGATC 186
Qy      181 ATCCGCCAGCTGCACCCCGCGCTGCAGAGCCGCGAGCGAGAGCTGAAGAGCTTGTCAAC 240
        |||
Db      187 ATCCGCCAGCTGCACCCCGCGCTGCAGAGCCGCGAGCGAGAGCTGAAGAGCTTGTCAAC 246
Qy      241 ACCGTGGCCACCTGTACTGTGGTGCACGAGAGATGAGGTCCGCGACACCAAGAGGCC 300
        |||
Db      247 ACCGTGGCCACCTGTACTGTGGTGCACGAGAGATGAGGTCCGCGACACCAAGAGGCC 306
Qy      301 CTGAGCAAGATCGAGAGAGCAGAACAGTGCACAGAGAGATCCAGCAGGCCGAGGCC 360
        |||
Db      307 CTGAGCAAGATCGAGAGAGCAGAACAGTGCACAGAGAGATCCAGCAGGCCGAGGCC 366
Qy      361 GCCGACAAAGGCAAGGTGAGGCCAGAACTACCCCATCTGTGCAAACTTGCAAGGCCAGATG 420
        |||
Db      367 GCCGACAAAGGCAAGGTGAGGCCAGAACTACCCCATCTGTGCAAACTTGCAAGGCCAGATG 426
Qy      421 GTGCAACAGGCCATCAGCCCCCGCACCCCTGAAAGCCTGGTGAAGGTGATGAGAGAAAG 480
        |||
Db      427 GTGCAACAGGCCATCAGCCCCCGCACCCCTGAAAGCCTGGTGAAGGTGATGAGAGAAAG 486
Qy      481 GCCTTCAGCCCCGAGGTGATCCCATGTTCAACCGCCTTGAGCGAGGGCGCCACCCCGCAG 540
        |||
Db      487 GCCTTCAGCCCCGAGGTGATCCCATGTTCAACCGCCTTGAGCGAGGGCGCCACCCCGCAG 546
Qy      541 GACCTGAACAAGATGTTGAACAACGTTGGCGCGCCACCAAGGCCGCTATGCAAGTCTGAAG 600
        |||
Db      547 GACCTGAACAAGATGTTGAACAACGTTGGCGCGCCACCAAGGCCGCTATGCAAGTCTGAAG 606
Qy      601 GACACCATCAACGAGAGAGCGCGCGAGTGGACCGCGTGCACCCCGTGCAAGCCGCGCCCC 660
        |||
Db      607 GACACCATCAACGAGAGAGCGCGCGAGTGGACCGCGTGCACCCCGTGCAAGCCGCGCCCC 666
Qy      661 ATGCGCCCGCGCGCAGATGCGGAGCGCCCGCGCGCAGCGACATCGCCGCGCACCAAGCACC 720
        |||
Db      667 ATGCGCCCGCGCGCAGATGCGGAGCGCCCGCGCGCAGCGACATCGCCGCGCACCAAGCACC 726
Qy      721 CTGCAAGAGCAGATCGCCTGAGTAGCAGCAACCCCGCATCCCGTGGCGCATCTTAC 780
        |||
Db      727 CTGCAAGAGCAGATCGCCTGAGTAGCAGCAACCCCGCATCCCGTGGCGCATCTTAC 786
Qy      781 AAGCGGTGATCATCTGCGCGCTGAACAAGATCGTGCAGATGTACAGCCCGTGAAGCATC 840
        |||
Db      787 AAGCGGTGATCATCTGCGCGCTGAACAAGATCGTGCAGATGTACAGCCCGTGAAGCATC 846
Qy      841 CTGACATCAAGCAGAGGCCCAAGAGGCCCTTCGCGACTACGTGAGCCGCTTCTTCAAG 900
        |||
Db      847 CTGACATCAAGCAGAGGCCCAAGAGGCCCTTCGCGACTACGTGAGCCGCTTCTTCAAG 906
Qy      901 ACCCTGCGCGCGCAGCAGAGCAACCCAGAGGTGAAGAACTGATGACCGACACCTGTGCTG 960
        |||
Db      907 ACCCTGCGCGCGCAGCAGAGCAACCCAGAGGTGAAGAACTGATGACCGACACCTGTGCTG 966
Qy      961 GTGCAAGAGCCCAACCCCGACTGCAAGACCATCTGCGCGCTCTGCGCCCGCGCGCAGC 1020
        |||
Db      967 GTGCAAGAGCCCAACCCCGACTGCAAGACCATCTGCGCGCTCTGCGCCCGCGCGCAGC 1026
Qy      1021 CTGAGAGAGATGATGACCGCTGCGCAGGCGTGGCGCGCCCGCAGCCACAAGGCCCGCGTG 1080
        |||
Db      1027 CTGAGAGAGATGATGACCGCTGCGCAGGCGTGGCGCGCCCGCAGCCACAAGGCCCGCGTG 1086
Qy      1081 CTGCGCGAGGCGATGAGCGCAGGCCAACACACAGCGTGTATGATGAGAGAGCAACTTCAAG 1140
        |||
Db      1087 CTGCGCGAGGCGATGAGCGCAGGCCAACACACAGCGTGTATGATGAGAGAGCAACTTCAAG 1146

```

```

Qy      1141 GGCCCCCGCGCATCTGTCAAGTGTCTTCAACTGCGGCAAGAGAGGCCACATCGCCCCCAAC 1200
        |||
Db      1147 GGCCCCCGCGCATCTGTCAAGTGTCTTCAACTGCGGCAAGAGAGGCCACATCGCCCCCAAC 1206
Qy      1201 TGCCGCGCCCCCGCAAGAGGCTGTGAAGTGCAGGCAAGAGAGGCCACACAGATGAAG 1260
        |||
Db      1207 TGCCGCGCCCCCGCAAGAGGCTGTGAAGTGCAGGCAAGAGAGGCCACACAGATGAAG 1266
Qy      1261 GACTGCACCGAGCGCGCAGGCCCACTTCTGGGCAAGATCTGGCCAGCCACAAGGGCCGC 1320
        |||
Db      1267 GACTGCACCGAGCGCGCAGGCCCACTTCTGGGCAAGATCTGGCCAGCCACAAGGGCCGC 1326
Qy      1321 CCGGCAACTTCTGTGACAGAGCCGCCGAGCCACCGCCCCCGCGGAGAGCTTCCGC 1380
        |||
Db      1327 CCGGCAACTTCTGTGACAGAGCCGCCGAGCCACCGCCCCCGCGGAGAGCTTCCGC 1386
Qy      1381 TTCGAGAGAGACCAACCCCGCGCAGAGCAGAGCAGAGACCGCGAGACCTTGACAGC 1440
        |||
Db      1387 TTCGAGAGAGACCAACCCCGCGCAGAGCAGAGCAGAGACCGCGAGACCTTGACAGC 1446
Qy      1441 CTGAAGAGCCTGTTCGCAACGACCCCTGAGCCAGTAA 1479
        |||
Db      1447 CTGAAGAGCCTGTTCGCAACGACCCCTGAGCCAGAAA 1485

```

```

RESULT 12
ACA03524
ID   ACA03524 standard; DNA; 2742 BP.
XX
AC   ACA03524;
XX
DT   22-MAY-2003 (first entry)
XX
DE   Synthetic DNA encoding immunogenic HIV peptide #7.
XX
KW   Immunogenic HIV polypeptide; human immunodeficiency virus; HIV; vaccine;
KW   gene therapy; packaging cell line; humoral immune response;
KW   cellular immune response; gene delivery vector; DNA immunisation; ds.
XX
OS   Synthetic.
XX
FN   WO2003004657-A1.
XX
PD   16-JAN-2003.
XX
PF   05-JUL-2002; 2002WO-US021421.
XX
PR   05-JUL-2001; 2001US-0303192P.
PR   31-AUG-2001; 2001US-0316860P.
PR   16-JAN-2002; 2002US-0349728P.
PR   16-JAN-2002; 2002US-0349793P.
PR   16-JAN-2002; 2002US-0349871P.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Zur Megede J, Barnett SW, Lian Y;
XX
WPI; 2003-221602/21.
XX
DR   WPI; 2003-221602/21.
XX
PT   New synthetic polynucleotides encoding antigenic HIV type B and/or type C
PT   polypeptides, useful as immunogenic compositions or vaccines for
PT   generating humoral or cellular immune responses against HIV in a subject,
PT   especially humans.
XX
PS   Example 1; Fig 12; 262pp; English.
XX
CC   The invention describes a synthetic polynucleotide encoding 2 or more
CC   immunogenic HIV polypeptides, where at least 2 of the polypeptides are
CC   derived from different HIV subtypes. The polynucleotide is useful for
CC   immunisation, generation of packaging cell lines, or production of HIV
CC   polypeptides. The polynucleotide and its encoded proteins are useful as
CC   immunogenic compositions or vaccines for generating humoral or cellular

```



PT New expression cassette comprising a polynucleotide sequence encoding a  
PT polypeptide including an HIV Gag, Env, Int, Nef, p15RaseH, Pol, Tat,  
PT Prot, or Rev polypeptide, useful for immunization, or generating  
PT packaging cell lines.

PS Disclosure; Fig 17; 301bp; English.

CC The invention relates to a novel expression cassette comprising a  
CC polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,  
CC Int, Nef, p15RaseH, Pol, Tat, Prot, or Rev polypeptide. The novel  
CC expression cassette can be used to treat HIV type C by gene therapy or  
CC used in the development of a vaccine. The gene delivery vector is  
CC administered intramuscularly, intramuscularly, intranasally,  
CC subcutaneously, intradermally, transdermally, intravaginally,  
CC intrarectally, orally or intravenously. The expression cassette is useful  
CC for immunisation, generating packaging cell lines and producing HIV  
CC polypeptides. This polynucleotide sequence represents the DNA of an HIV  
CC Type C related sequence of the invention.

XX Sequence 2742 BP; 578 A; 1020 C; 859 G; 285 T; 0 U; 0 Other;

Query Match 99.8%; Score 1475.8; DB 9; Length 2742;

Best Local Similarity 99.9%; Pred. No. 5.4e-185;

Matches 1477; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 ATGGGGCGCGCGCCAGCATCTCTGGCGGGCGGCAAGCTGGAACGCTTGGGAGCGCATCCGC 60
DB 7 ATGGGGCGCGCGCCAGCATCTCTGGCGGGCGGCAAGCTGGAACGCTTGGGAGCGCATCCGC 66
QY 61 CTGCGCGCGCGCGGCAAGAGTCTTACATGATGAAGCACTGTGTGGGGCCAGCGCGAG 120
DB 67 CTGCGCGCGCGCGGCAAGAGTCTTACATGATGAAGCACTGTGTGGGGCCAGCGCGAG 126
QY 121 CTGGAGAAGTTCGCTCTGAACCGCGCGCTGTGTGAGACCAAGGCGCTGCAAGAGATC 180
DB 127 CTGGAGAAGTTCGCTCTGAACCGCGCGCTGTGTGAGACCAAGGCGCTGCAAGAGATC 186
QY 181 ATCCCGCAGCTGACCGCGCGCTGTGAGACCGGCAAGGAGCTGAAGAGCTTTCAC 240
DB 187 ATCCCGCAGCTGACCGCGCGCTGTGAGACCGGCAAGGAGCTGAAGAGCTTTCAC 246
QY 241 ACCGTGGCCACCTCTGTACTGCGTGCACGAGAGATCGAGTCCGCGACCAAGAGAGCC 300
DB 247 ACCGTGGCCACCTCTGTACTGCGTGCACGAGAGATCGAGTCCGCGACCAAGAGAGCC 306
QY 301 CTGGACAAGATCGAGAGAGAGCAACAAGTCCAGCAAGATTCAGAGCGCGAGGCC 360
DB 307 CTGGACAAGATCGAGAGAGAGCAACAAGTCCAGCAAGATTCAGAGCGCGAGGCC 366
QY 361 GCCGACAAGGCAAGGTGAGGCAACTACCCCATCTGTGCAAACTTGCAGGGCGCATG 420
DB 367 GCCGACAAGGCAAGGTGAGGCAACTACCCCATCTGTGCAAACTTGCAGGGCGCATG 426
QY 421 GTGCACCAAGGCGCATCAGCGCGCGCAACCTGAAAGCTGTGAAGTATCGAGAGAG 480
DB 427 GTGCACCAAGGCGCATCAGCGCGCGCAACCTGAAAGCTGTGAAGTATCGAGAGAG 486
QY 481 GCCTTCAGCGCGGAGGTGATCCCATGTTTACCGCGCTGAGCGAGGGCGCCACCC 540
DB 487 GCCTTCAGCGCGGAGGTGATCCCATGTTTACCGCGCTGAGCGAGGGCGCCACCC 546
QY 541 GACCTGAACACGATGTTGAACACCGTGGCGCGCAACAGCGCGCCATGCAAGTCTGAAG 600
DB 547 GACCTGAACACGATGTTGAACACCGTGGCGCGCAACAGCGCGCCATGCAAGTCTGAAG 606
QY 601 GACACCATCAACGAGAGCGCGCGGAGTGGGACCGCGTGCACCGCGTGCACCGCGGCC 660
DB 607 GACACCATCAACGAGAGCGCGCGGAGTGGGACCGCGTGCACCGCGTGCACCGCGGCC 666
QY 661 ATGCGCGCGCGCGCAGATGCGCGAGCGCGCGCGCAGCGACATCGCGCGGCAACCA 720
DB 667 ATGCGCGCGCGCGCAGATGCGCGAGCGCGCGCGCAGCGACATCGCGCGGCAACCA 726
```

```
QY 721 CTGCAGAGCAGATCGCTGTGATGACCAACCCCCCATCCCGTGGGCGCATCTTAC 780
DB 727 CTGCAGAGCAGATCGCTGTGATGACCAACCCCCCATCCCGTGGGCGCATCTTAC 786
QY 781 AAGCGGTGATCATCTGGGCGCTGAACAAGATCGTGGGATGTACAGCGCGTAGCATC 840
DB 787 AAGCGGTGATCATCTGGGCGCTGAACAAGATCGTGGGATGTACAGCGCGTAGCATC 846
QY 841 CTGCACATCAAGCAGGCGCGCAAGAGCGCTTCCGCACTACGTGACCGCTTCTTCAAG 900
DB 847 CTGCACATCAAGCAGGCGCGCAAGAGCGCTTCCGCACTACGTGACCGCTTCTTCAAG 906
QY 901 ACCCTGCGCGCGAGCAGACCAAGAGGTGAAGAACTGATGACCGACACCTGTCTG 960
DB 907 ACCCTGCGCGCGAGCAGACCAAGAGGTGAAGAACTGATGACCGACACCTGTCTG 966
QY 961 GTGCAGAACGCAACCGCGACTGTGCAAGACCATCTGCGCGCTTGGCGCGCGCGCAGC 1020
DB 967 GTGCAGAACGCAACCGCGACTGTGCAAGACCATCTGCGCGCTTGGCGCGCGCGCAGC 1026
QY 1021 CTGCAGAGATGATGACCGCTGTGCAAGGCGCTGCGCGCGCGCGCAGCAAGCGCGCTG 1080
DB 1027 CTGCAGAGATGATGACCGCTGTGCAAGGCGCTGCGCGCGCGCGCAGCAAGCGCGCTG 1086
QY 1081 CTGCGGAGGCGATGAGCCAGGCGCAACCAAGCGTATGATGACAGAGCAACTTCAAG 1140
DB 1087 CTGCGGAGGCGATGAGCCAGGCGCAACCAAGCGTATGATGACAGAGCAACTTCAAG 1146
QY 1141 GCGCGCGCGCGCATCTGCAAGTGTCTTCACTGCGCGCAAGAGGCGCATGCGCGCAAC 1200
DB 1147 GCGCGCGCGCGCATCTGCAAGTGTCTTCACTGCGCGCAAGAGGCGCATGCGCGCAAC 1206
QY 1201 TGCGCGCGCGCGCAAGAGGCTGTGGAAGTGCAGGCGCAAGGAGCCACAGATGAAG 1260
DB 1207 TGCGCGCGCGCGCAAGAGGCTGTGGAAGTGCAGGCGCAAGGAGCCACAGATGAAG 1266
QY 1261 GACTGACCGGAGCGCGCAACTTCTGTGGCAAGATCTGCGCGCAAGGCGCGC 1320
DB 1267 GACTGACCGGAGCGCGCAACTTCTGTGGCAAGATCTGCGCGCAAGGCGCGC 1326
QY 1321 CCGGCAACTTCTGTGAGAGCGCGCGCGCAACCGCGCGCGCGCGAGAGCTTCCGC 1380
DB 1327 CCGGCAACTTCTGTGAGAGCGCGCGCGCGCAACCGCGCGCGCGCGAGAGCTTCCGC 1386
QY 1381 TTGAGAGAGACCAACCGCGCGCAAGAGCAAGAGCAAGAGCGCGAGACCTGACCAAG 1440
DB 1387 TTGAGAGAGACCAACCGCGCGCAAGAGCAAGAGCAAGAGCGCGAGACCTGACCAAG 1446
QY 1441 CTGAAGAGCTGTTGGCAAGCAACCGCTGAGCGCACTAA 1479
DB 1447 CTGAAGAGCTGTTGGCAAGCAACCGCTGAGCGCACTAA 1485
```

#### RESULT 14

AD13230  
ID AD13230 standard; DNA; 3930 BP.

XX AD13230;

DT 18-DEC-2003 (first entry)

DB DNA of HIV construct GagComp[Polmut\_C SEQ ID NO 9.

XX expression cassette; HIV Gag; Env; Int; Nef; p15RaseH; Pol; Tat; Prot;

XX Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; ds.

OS Human immunodeficiency virus.

PN WO200304620-A2.

PD 16-JAN-2003.

PF 05-JUL-2002; 2002WO-US021420.

XX 05-JUL-2001; 2001US-0303192P.  
PR 31-AUG-2001; 2001US-0316860P.  
PR 16-JAN-2002; 2002US-0349871P.

XX (CHIR ) CHIRON CORP.  
PA (UYST-) UNIV STRLEMBOSCH.

XX Zur Megede J, Barnett SW, Lian Y, Engelbrecht S, Van Rensburg EJ;  
PI  
XX WPI; 2003-221593/21.

XX New expression cassette comprising a polynucleotide sequence encoding a  
PT polypeptide including an HIV Gag, Env, Int, Nef, p15NaseH, Pol, Tat,  
PT Prot, or Rev polypeptide, useful for immunization, or generating  
PT packaging cell lines.

XX Disclosure; Fig 6; 301pp; English.

XX The invention relates to a novel expression cassette comprising a  
CC polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,  
CC Int, Nef, p15NaseH, Pol, Tat, Prot, or Rev polypeptide. The novel  
CC expression cassette can be used to treat HIV type C by gene therapy or  
CC used in the development of a vaccine. The gene delivery vector is  
CC administered intramuscularly, intramuscularly, intravenously,  
CC subcutaneously, intradermally, transdermally, intravaginally,  
CC intrarectally, orally or intravenously. The expression cassette is useful  
CC for immunisation, generating packaging cell lines and producing HIV  
CC polypeptides. This polynucleotide sequence represents the DNA of an HIV  
CC Type C related sequence of the invention.

XX SQ Sequence 3930 BP; 890 A; 1365 C; 1214 G; 461 T; 0 U; 0 Other;

Query Match 99.8%; Score 1475.8; DB 9; Length 3930;

Best Local Similarity 99.9%; Pred. No. 5.1e-185;  
Matches 1477; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGGCGCCCGCGCAGCATCTCGCGCGCGGCAAGCTGACGCTGGAGCGCATTCGCG 60  
DB 7 ATGGGCGCCCGCGCAGCATCTCGCGCGCGGCAAGCTGACGCTGGAGCGCATTCGCG 66  
QY 61 CTGCGCGCCCGCGGCAAGAGTGTATCATGTATGAAGCACTGTGTGGGCGAGCGCGAG 120  
DB 67 CTGCGCGCCCGCGGCAAGAGTGTATCATGTATGAAGCACTGTGTGGGCGAGCGCGAG 126  
QY 121 CTGGAAGAAGTTCGCTCTGAACCCCGGCTGTGAGACCAAGGCGCTGCAAGCATC 180  
DB 127 CTGGAAGAAGTTCGCTCTGAACCCCGGCTGTGAGACCAAGGCGCTGCAAGCATC 186  
QY 181 ATCCGCGAGCTGACCCCGGCTGTGAGACCGGCGAGGAGCTGAAGGCTGTTCAAC 240  
DB 187 ATCCGCGAGCTGACCCCGGCTGTGAGACCGGCGAGGAGCTGAAGGCTGTTCAAC 246  
QY 241 ACCGTGGCCACCTGTACTGCTGTGACGAGAGATGAGTCCGCAACCAAGAGGCGC 300  
DB 247 ACCGTGGCCACCTGTACTGCTGTGACGAGAGATGAGTCCGCAACCAAGAGGCGC 306  
QY 301 CTGGAAGAAGTTCGAGAGGAGCAAGTGTGCAAGAGATCAAGGCGCGAGGCGC 360  
DB 307 CTGGAAGAAGTTCGAGAGGAGCAAGTGTGCAAGAGATCAAGGCGCGAGGCGC 366  
QY 361 GCCGACAAGGCGAAGTGTAGCCAGACTACCCCATGTGCAAGACTGAGGCGCGAGATG 420  
DB 367 GCCGACAAGGCGAAGTGTAGCCAGACTACCCCATGTGCAAGACTGAGGCGCGAGATG 426  
QY 421 GTGCAACAGGCGATCAAGCCCGGCAAGCTGAAAGCTGTGAAAGTGTGAGAGAG 480  
DB 427 GTGCAACAGGCGATCAAGCCCGGCAAGCTGAAAGCTGTGAAAGTGTGAGAGAG 486  
QY 481 GCTTTCAGCCCGGAGGTATCCCAATGTTCACCGCTGTAGCGAGGCGCGACCCCGAG 540  
DB 487 GCTTTCAGCCCGGAGGTATCCCAATGTTCACCGCTGTAGCGAGGCGCGACCCCGAG 546

QY 541 GACCTGAACACGATGTTGAACACCGTGGCGGCGCACCAAGGCGCGCATGAGATGCTGAAG 600  
DB 547 GACCTGAACACGATGTTGAACACCGTGGCGGCGCACCAAGGCGCGCATGAGATGCTGAAG 606  
QY 601 GACACCATCAACGAGAGAGGCGCGCGAGTGGACCGGCTGACACCGCGCGCGCGC 660  
DB 607 GACACCATCAACGAGAGAGGCGCGCGAGTGGACCGGCTGACACCGCGCGCGCGC 666  
QY 661 ATGCGCCCGCGGCGAGATGCGCGGAGCGCGCGCGCGCGAGATGCGCGCGCGCGAGC 720  
DB 667 ATGCGCCCGCGGCGAGATGCGCGGAGCGCGCGCGCGCGAGATGCGCGCGCGAGC 726  
QY 721 CTGCAAGAGCAGATGCGCTGTGATGACCAAGCAACCCCGCATCCCGCTGGCGGACATCTAC 780  
DB 727 CTGCAAGAGCAGATGCGCTGTGATGACCAAGCAACCCCGCATCCCGCTGGCGGACATCTAC 786  
QY 781 AAGCGGTGATCATCTGCGGCTGAAACAGATGCTGCGGATGTAACGCGCGTGAGCATC 840  
DB 787 AAGCGGTGATCATCTGCGGCTGAAACAGATGCTGCGGATGTAACGCGCGTGAGCATC 846  
QY 841 CTGCAATCAAGCAGGCG 900  
DB 847 CTGCAATCAAGCAGGCG 906  
QY 901 ACCCTGCG 960  
DB 907 ACCCTGCG 966  
QY 961 GTGCAAGACGCAACCCCGACTGCAAGACCATCTGCGGCTGCGCGCGCGCGCGCGCGCG 1020  
DB 967 GTGCAAGACGCAACCCCGACTGCAAGACCATCTGCGGCTGCGCGCGCGCGCGCGCGCG 1026  
QY 1021 CTGGAAGAGATGATGACCGCTGTGCAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGTG 1080  
DB 1027 CTGGAAGAGATGATGACCGCTGTGCAAGGCGCGCGCGCGCGCGCGCGCGCGCGTG 1086  
QY 1081 CTGCGCGAGGCGAGTGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1140  
DB 1087 CTGCGCGAGGCGAGTGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1146  
QY 1141 GCG 1200  
DB 1147 GCG 1206  
QY 1201 TGC CG 1260  
DB 1207 TGC CG 1266  
QY 1261 GACTGACCGAGGCG 1320  
DB 1267 GACTGACCGAGGCG 1326  
QY 1321 CCGCGCAACTTCTGCAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380  
DB 1327 CCGCGCAACTTCTGCAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1386  
QY 1381 TTGAGAGAGACCAACCG 1440  
DB 1387 TTGAGAGAGACCAACCG 1446  
QY 1441 CTGAAGAGCGCTGTTGCGCAACGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1479  
DB 1447 CTGAAGAGCGCTGTTGCGCAACGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1485

RESULT 15

ADCI3231  
ID ADCI3231 standard; DNA; 3930 BP.

XX ADCI3231;

XX 18-DEC-2003 (first entry)



DE DNA of HIV construct GagComp1PolmutAlt\_C SEQ ID NO 10.

KX expression cassette; HIV Gag; Env; Int; Nef; p15RnaseH; Pol; Tat; Prot;  
KM Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; de.  
XX  
OS Human immunodeficiency virus.

XX WO2003004620-A2.

XX 16-JAN-2003.

XX 05-JUL-2002; 2002WO-US021420.

XX 05-JUL-2001; 2001US-0303192P.

XX 31-AUG-2001; 2001US-0316860P.

XX 16-JAN-2002; 2002US-0349871P.

XX (CHIR ) CHIRON CORP.

XX (UYST-) UNIV STELLENBOSCH.

XX Zur Megele J, Barnett SW, Lian Y, Engelbrecht S, Van Rensburg BJ;

XX WPI; 2003-221593/21.

XX New expression cassette comprising a polynucleotide sequence encoding a

XX polypeptide including an HIV Gag, Env, Int, Nef, p15RnaseH, Pol, Tat,

XX Prot, or Rev polypeptide, useful for immunization, or generating

XX packaging cell lines.

XX Disclosure; Fig 7; 301pp; English.

XX The invention relates to a novel expression cassette comprising a

XX polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,

XX Int, Nef, p15RnaseH, Pol, Tat, Prot, or Rev polypeptide. The novel

XX expression cassette can be used to treat HIV type C by gene therapy or

XX used in the development of a vaccine. The gene delivery vector is

XX administered intramuscularly, intramuscularly, intranasally,

XX subcutaneously, intradermally, transdermally, intravaginally,

XX intrarectally, orally or intravenously. The expression cassette is useful

XX for immunisation, generating packaging cell lines and producing HIV

XX polypeptides. This polynucleotide sequence represents the DNA of an HIV

XX Type C related sequence of the invention.

XX Sequence 3930 BP; 889 A; 1365 C; 1214 G; 462 T; 0 U; 0 Other;

XX Query Match 99.8%; Score 1475.8; DB 9; Length 3930;

XX Best Local Similarity 99.9%; Pred. No. 5.1e-185;

XX Matches 1477; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGGCGCCCGCGCAGCATCTCGCGCGCGCAAGCTGAGCGCTGGAGCGCATCCGC 60

QY 7 ATGGGCGCCCGCGCAGCATCTCGCGCGCGCAAGCTGAGCGCTGGAGCGCATCCGC 66

QY 61 CTGCGCCCGCGCGCAAGAGTGTACATGATGAAGCACTGTGTGGCCAGCCGCGAG 120

QY 67 CTGCGCCCGCGCGCAAGAGTGTACATGATGAAGCACTGTGTGGCCAGCCGCGAG 126

QY 121 CTGAGAGAGTTGCGCTGAACCCCGCGCTGCTGAGAGCAAGAGGCTGCAAGAGATC 180

QY 127 CTGAGAGAGTTGCGCTGAACCCCGCGCTGCTGAGAGCAAGAGGCTGCAAGAGATC 186

QY 181 ATCCGCGAGCTGCAACCCCGCGCTGCAAGAGGCGAGAGCTGAAGAGCTGTTCAAC 240

QY 187 ATCCGCGAGCTGCAACCCCGCGCTGCAAGAGGCGAGAGCTGAAGAGCTGTTCAAC 246

QY 241 ACCGTGCGCAACCTGTATCTGCTGCAAGAGAGATCGAGTCCGCGACCAAGAGGCC 300

QY 247 ACCGTGCGCAACCTGTATCTGCTGCAAGAGAGATCGAGTCCGCGACCAAGAGGCC 306

QY 301 CTGAGACAGATCGAG 360

QY 307 CTGAGACAGATCGAG 366

QY 361 GCCGACAAGGGCAAGGTGAGCCAGAACTACCCCATCTGTGAGAACTGACAGGCCAGATG 420

QY 367 GCCGACAAGGGCAAGGTGAGCCAGAACTACCCCATCTGTGAGAACTGACAGGCCAGATG 426

QY 421 GTGCAACAGGGCATCAGCCCGCGCAACCTGAAGCCCTGGTGAAGGTGATCGAGAGAG 480

QY 427 GTGCAACAGGGCATCAGCCCGCGCAACCTGAAGCCCTGGTGAAGGTGATCGAGAGAG 486

QY 481 GCCTTCAGCCCGCGAGGTGATCCCATATGTTCAACCGCCCTGAGCGAGGGCGCCACCC 540

QY 487 GCCTTCAGCCCGCGAGGTGATCCCATATGTTCAACCGCCCTGAGCGAGGGCGCCACCC 546

QY 541 GACCTGAACACGATGTTGAACACCGTGGCGCGCCACCAAGGCCCATGAGATCTGAAG 600

QY 547 GACCTGAACACGATGTTGAACACCGTGGCGCGCCACCAAGGCCCATGAGATCTGAAG 606

QY 601 GACACCATCAACGAGAGAGGGCGCGAGTGGAGCCGCGTGCACCCCGTGCACCGCGGCC 660

QY 607 GACACCATCAACGAGAGAGGGCGCGAGTGGAGCCGCGTGCACCCCGTGCACCGCGGCC 666

QY 661 ATCCGCCCCCGCGCAAGATGCGCGAGCCCGCGCGAGCGACATGCGCGCGCACCAAGCAC 720

QY 667 ATCCGCCCCCGCGCAAGATGCGCGAGCCCGCGCGAGCGACATGCGCGCGCACCAAGCAC 726

QY 721 CTGCAAGAGAGATGCGCTGATGACCAAGCAACCCCATCTGCGCGCGCATCTAC 780

QY 727 CTGCAAGAGAGATGCGCTGATGACCAAGCAACCCCATCTGCGCGCGCATCTAC 786

QY 781 AAGCGGTGATCATCTCGCGCTGAACAAGATGTCGAGATGTAAGCCCGTGAACATC 840

QY 787 AAGCGGTGATCATCTCGCGCTGAACAAGATGTCGAGATGTAAGCCCGTGAACATC 846

QY 841 CTGACATCAAGAGAGGGCGCGCAAGAGCCCTTCCGCACTAGTGAGACCGCTTCTCAAG 900

QY 847 CTGACATCAAGAGAGGGCGCGCAAGAGCCCTTCCGCACTAGTGAGACCGCTTCTCAAG 906

QY 901 ACCCTGCGCGCGAG 960

QY 907 ACCCTGCGCGCGAG 966

QY 961 GTGCAAG 1020

QY 967 GTGCAAG 1026

QY 1021 CTGAGAGAGATGATGACCGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080

QY 1027 CTGAGAGAGATGATGACCGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1086

QY 1081 CTGCGCGAGCGATGAG 1140

QY 1087 CTGCGCGAGCGATGAG 1146

QY 1141 GGGCCCCCGCGCATGCTGAAGTCTTCACTGCGGCAAGAGAGAGAGAGAGAGAGAGAGAG 1200

QY 1147 GGGCCCCCGCGCATGCTGAAGTCTTCACTGCGGCAAGAGAGAGAGAGAGAGAGAGAGAG 1206

QY 1201 TGCGCGCGCCCCCGCAAG 1260

QY 1207 TGCGCGCGCCCCCGCAAG 1266

QY 1261 GACTGCAAG 1320

QY 1267 GACTGCAAG 1326

QY 1321 CCGGCAACTTCTGAG 1380

QY 1327 CCGGCAACTTCTGAG 1386

QY 1381 TTGAG 1440

QY 1387 TTGAG 1446

QY 1441 CTGAAGAGCTGTTGCGCAAG 1479



db 1447 CTGAAGCCTGTTCGCAAGACCCCTGAGCCAGAA 1485

Search completed: May 28, 2004, 09:49:09  
Job time : 689.39 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 28, 2004, 09:05:41 ; Search time 6077.87 Seconds  
(without alignments)  
10547.179 Million cell updates/sec

Title: US-09-475-704A-3

Perfect score: 1479  
Sequence: 1 atggcgccgcgcgcagcat.....acgacccctgagccagtaa 1479

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 segs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenBankl: \*  
1: gb\_ba: \*  
2: gb\_htg: \*  
3: gb\_in: \*  
4: gb\_om: \*  
5: gb\_ov: \*  
6: gb\_pat: \*  
7: gb\_ph: \*  
8: gb\_pl: \*  
9: gb\_pr: \*  
10: gb\_ro: \*  
11: gb\_sts: \*  
12: gb\_sy: \*  
13: gb\_un: \*  
14: gb\_vi: \*  
15: em\_ba: \*  
16: em\_fun: \*  
17: em\_hum: \*  
18: em\_in: \*  
19: em\_mu: \*  
20: em\_om: \*  
21: em\_or: \*  
22: em\_ov: \*  
23: em\_pat: \*  
24: em\_ph: \*  
25: em\_pl: \*  
26: em\_ro: \*  
27: em\_sts: \*  
28: em\_un: \*  
29: em\_vi: \*  
30: em\_htg\_hum: \*  
31: em\_htg\_inv: \*  
32: em\_htg\_other: \*  
33: em\_htg\_mus: \*  
34: em\_htg\_pln: \*  
35: em\_htg\_rod: \*  
36: em\_htg\_mam: \*  
37: em\_htg\_vrt: \*  
38: em\_by: \*  
39: em\_htgo\_hum: \*  
40: em\_htgo\_mus: \*  
41: em\_htgo\_other: \*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1479	100.0	1479	6	AX455887 Sequence
2	1479	100.0	1479	6	AX468543 Sequence
3	1473.4	99.6	1479	6	AX468547 Sequence
4	1463	98.9	1479	6	AX455904 Sequence
5	1328.2	89.8	1485	12	AY181195 Synthetic
6	1288.8	87.1	1491	6	AX455983 Sequence
7	1288	87.1	1494	6	AX455935 Sequence
8	1276.8	86.3	1509	6	AX455888 Sequence
9	1275.2	86.2	1509	6	AX468544 Sequence
10	1273.6	86.1	1509	6	AX468548 Sequence
11	1260.8	85.2	1509	6	AX455905 Sequence
12	1221.2	82.6	9166	6	AX427930 Sequence
13	1206	81.5	9788	6	AX427936 Sequence
14	1205.4	81.5	4288	6	AX149648 Sequence
15	1197.8	81.0	1509	12	AF201927 Synthetic
16	1197.8	81.0	1515	6	BD263637 Improved
17	1197.8	81.0	1515	6	AR373320 Sequence
18	1197.8	81.0	4472	6	BD263697 Improved
19	1197.8	81.0	4472	6	AR373380 Sequence
20	1197.8	81.0	4608	6	BD263698 Improved
21	1197.8	81.0	4608	6	AR373381 Sequence
22	1197.8	81.0	4689	6	BD263696 Improved
23	1197.8	81.0	4689	6	AR373379 Sequence
24	1197.8	81.0	4766	6	BD263695 Improved
25	1197.8	81.0	4766	6	AR373378 Sequence
26	1195.8	80.9	2031	6	BD263640 Improved
27	1195.8	80.9	2031	6	AR373323 Sequence
28	1159.6	78.4	1548	12	AF287354 Synthetic
29	1134.4	76.7	1539	6	AX306431 Sequence
30	1131.8	76.5	6438	6	AX427919 Sequence
31	1131.8	76.5	9194	6	AX427925 Sequence
32	1131.8	76.5	9194	6	AX427926 Sequence
33	1131.8	76.5	12411	6	AX427927 Sequence
34	1120.4	75.8	9169	6	AX427931 Sequence
35	1116.6	75.5	9170	6	AX427928 Sequence
36	1116.6	75.5	9189	6	AX427921 Sequence
37	1116.6	75.5	9792	6	AX427932 Sequence
38	1114.6	75.4	1847	12	AF202464 Synthetic
39	1114.6	75.4	1847	12	AF202465 Synthetic
40	1114.6	75.4	1853	6	BD263638 Improved
41	1114.6	75.4	1853	6	AR373321 Sequence
42	1114.6	75.4	1865	6	BD263700 Improved
43	1114.6	75.4	1865	6	BD263701 Improved
44	1114.6	75.4	1865	6	AR373383 Sequence
45	1114.6	75.4	1865	6	AR373384 Sequence

ALIGNMENTS

RESULT 1  
AX455887 1479 bp DNA linear PAT 06-JUL-2002  
LOCUS AX455887 Sequence 3 from Patent WO0204493.  
DEFINITION AX455887  
ACCESSION AX455887  
VERSION AX455887.1 GI:21714880  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
artificial sequences.  
REFERENCE 1 zur Megeide, J., Barnett, S.W., Engelbrecht, S. and van Rensburg, E.  
AUTHORS Polynucleotides encoding antigenic hiv type c polypeptides,  
TITLB polypeptides and uses thereof  
JOURNAL Patent: WO 0204493-A 3 17-JAN-2002;

Pred. No. is the number of results predicted by chance to have a



Oy	61	CTGCGCCCCGGCGGCAAGAAAGTGTCTACATGATGAAGCACTGTGTGTGGCCAGCGCGAG	120
Db	61	CTGCGCCCCGGCGGCAAGAAAGTGTCTACATGATGAAGCACTGTGTGTGGCCAGCGCGAG	120
Oy	121	CTGGAAGAAGTTCGCTCTGAACCCCGGCTGTGTGAAGCACGAGAGGCTGCAAGCAGATC	180
Db	121	CTGGAAGAAGTTCGCTCTGAACCCCGGCTGTGTGAAGCACGAGAGGCTGCAAGCAGATC	180
Oy	181	ATCCGCCAGCTGCACCCCGCTGCAGACCGGACGAGAGCTGAAGCCTGTTCAAC	240
Db	181	ATCCGCCAGCTGCACCCCGCTGCAGACCGGACGAGAGCTGAAGCCTGTTCAAC	240
Oy	241	ACCGTGGCCACCCTGTACTGCGTGCACGAGAAGATCGAGTCCGCGACACCAAGAGGCC	300
Db	241	ACCGTGGCCACCCTGTACTGCGTGCACGAGAAGATCGAGTCCGCGACACCAAGAGGCC	300
Oy	301	CTGGAACAAGATCGAGAGGAGCAGAACAAAGTGCACGACAAGATCCAGCAGGCCGAGGCC	360
Db	301	CTGGAACAAGATCGAGAGGAGCAGAACAAAGTGCACGACAAGATCCAGCAGGCCGAGGCC	360
Oy	361	GCCGCAAGGGCAAGGTGAGCCAGAACTACCCCATCTGTGCAGAACTTGACAGGGCCAGATG	420
Db	361	GCCGCAAGGGCAAGGTGAGCCAGAACTACCCCATCTGTGCAGAACTTGACAGGGCCAGATG	420
Oy	421	GTGCACCAAGGCCATCAGCCCCCGCACCTGAAACGCTGGGTGAAGGTGATGAGAGAAAG	480
Db	421	GTGCACCAAGGCCATCAGCCCCCGCACCTGAAACGCTGGGTGAAGGTGATGAGAGAAAG	480
Oy	481	GCCTTCAGCCCCCGAGGTGATCCCACTGTTCACCGCCTTGAGCGAGGGCGCCACCCCCAG	540
Db	481	GCCTTCAGCCCCCGAGGTGATCCCACTGTTCACCGCCTTGAGCGAGGGCGCCACCCCCAG	540
Oy	541	GACCTGAACAACGATGTTGAACAACGTGGGGCGGCGCACAGGCCGCATGCAAGATCTGAAG	600
Db	541	GACCTGAACAACGATGTTGAACAACGTGGGGCGGCGCACAGGCCGCATGCAAGATCTGAAG	600
Oy	601	GACACCATCAACGAGAGGGCGCGCGAGTGGGACCGCGTGCACCCCGTGCACGCGGCCCC	660
Db	601	GACACCATCAACGAGAGGGCGCGCGAGTGGGACCGCGTGCACCCCGTGCACGCGGCCCC	660
Oy	661	ATCGCCCCCGCCAGATGCGCGAGCCCCCGGCAAGCAATCGCCGGCACCAACGACACC	720
Db	661	ATCGCCCCCGCCAGATGCGCGAGCCCCCGGCAAGCAATCGCCGGCACCAACGACACC	720
Oy	721	CTGCAGAGCAGATCGCTTGATGACCAAGCAACCCCTCATCCCTGTGGCGCATCTAC	780
Db	721	CTGCAGAGCAGATCGCTTGATGACCAAGCAACCCCTCATCCCTGTGGCGCATCTAC	780
Oy	781	AAGCGGTGATCATCTCTGGGCTTGAAACAAGATCTGTGGATGTACAGCCCCGTGAGCATC	840
Db	781	AAGCGGTGATCATCTCTGGGCTTGAAACAAGATCTGTGGATGTACAGCCCCGTGAGCATC	840
Oy	841	CTGGAATCAAGCAGGGGCCCAAGAGACCCTTCCGCGACTACGTGAACCGCTTCTTCAAG	900
Db	841	CTGGAATCAAGCAGGGGCCCAAGAGACCCTTCCGCGACTACGTGAACCGCTTCTTCAAG	900
Oy	901	ACCCTGCGCGCCGAGCAGAGCACCCAGAGAGTGAAGAACTGGATGACCGACACCTGCTG	960
Db	901	ACCCTGCGCGCCGAGCAGAGCACCCAGAGAGTGAAGAACTGGATGACCGACACCTGCTG	960
Oy	961	GTGCAGAACGCCAACCCCGACTGCAGAACCATCTGCGCGCTTCTCGGCCCCCGGGCCAGC	1020
Db	961	GTGCAGAACGCCAACCCCGACTGCAGAACCATCTGCGCGCTTCTCGGCCCCCGGGCCAGC	1020
Oy	1021	CTGGAAGAGATGATGACCGCTGTGCAGGGCGTGGGGCGGCCACGCAACAAGGCCCGGTG	1080
Db	1021	CTGGAAGAGATGATGACCGCTGTGCAGGGCGTGGGGCGGCCACGCAACAAGGCCCGGTG	1080
Oy	1081	CTGGCCGAGGGCGATGAGCCAGGCGCAACACCAAGCCTGATGATGCAAGAAAGCACTTCAAG	1140
Db	1081	CTGGCCGAGGGCGATGAGCCAGGCGCAACACCAAGCCTGATGATGCAAGAAAGCACTTCAAG	1140
Oy	1141	GCGCCCCGGCGCATGCTCAAGTCTTCAACTGCGCGCAAGAGGCCCATATCGCCCGCAAC	1200

Db	1141	GGCCCCGGCGCATCGTCAAGTGTCTCAACTGCGGCAAGAGGGGCACATGCGCCGCAAC	1200
QY	1201	TGCGCGCCCCCGCGAAGAGGCGTGTCTGGAAGTGCGGCAAGAGGGCCACAGATGAAG	1260
Db	1201	TGCGCGCGCCCCCGCGAAGAGGCGTGTCTGGAAGTGCGGCAAGAGGGCCACAGATGAAG	1260
QY	1261	GACTGCACCGAGCGGCCCAACTTCTGTGGCAAGATCTGSCCCAGCCACAAGGSCCGC	1320
Db	1261	GACTGCACCGAGCGGCCCAACTTCTGTGGCAAGATCTGSCCCAGCCACAAGGSCCGC	1320
QY	1321	CCCGGCAACTTCTGTGAGAGCCCGCCCGAGGCCACCGCCCCCGCGAGAGCTTCCGC	1380
Db	1321	CCCGGCAACTTCTGTGAGAGCCCGCCCGAGGCCACCGCCCCCGCGAGAGCTTCCGC	1380
QY	1381	TTGAGAGAGACCAACCCCGCGCCAGAGGCAAGAGCAAGACCTCGAGACCAGC	1440
Db	1381	TTGAGAGAGACCAACCCCGCGCCAGAGGCAAGAGCAAGACCTCGAGACCAGC	1440
QY	1441	CTGAAGAGCCTGTTCGGCAACGACCCCTGAGCCAGTAA	1479
Db	1441	CTGAAGAGCCTGTTCGGCAACGACCCCTGAGCCAGTAA	1479

RESULT 3	AX468547	1479 bp	DNA	linear	PAT 16-JUL-2002
LOCUS	AX468547	Sequence 67 from Patent WO0226209.			
DEFINITION	AX468547				
ACCESSION	AX468547				
VERSION	AX468547.1	GI:21901377			
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
FEATURES					
source					
ORIGIN					
Query Match	99.6%;	Score 1473.4;	DB 6;	length 1479;	
Best Local Similarity	99.3%;	Pred. No. 1.6e-150;			
Matches 1468;	Conservative 10;	Mismatches 1;	Indels 0;	Gaps 0;	
QY	1 ATGGCGCCCGCGCGCCAGCATCTCTGCGCGCGCGGCAAGCTGGACGCTTGGAGCGCATCCGC				60
DB	1 ATGGCGCCCGCGCGCCAGCATCTCTGCGCGCGCGGCAAGCTGGACGCTTGGAGCGCATCCGC				60
QY	61 CTGGCGCCCGCGCGCGCAAGAGTGCTACATGATGAAAGCACCTGTGTGGGCCAGCGCGAG				120
DB	61 CTGGCGCCCGCGCGCGCAAGAGTGCTACATGATGAAAGCACCTGTGTGGGCCAGCGCGAG				120
QY	121 CTGAGAGAAGTTGCGCCCTGAACCCCGGCTGTGAGAGACGAGCGGCTGCAAGCAGATC				180
DB	121 CTGAGAGAAGTTGCGCCCTGAACCCCGGCTGTGAGAGACGAGCGGCTGCAAGCAGATC				180
QY	181 ATCCGCCAGCTGCACTCCCGCCCTGAGACCGGCAAGGAGCTGAAAGCCTGTTCAAC				240
DB	181 ATCCGCCAGCTGCACTCCCGCCCTGAGACCGGCAAGGAGCTGAAAGCCTGTTCAAC				240
QY	241 ACCGTGGCCACCCCTGTACTGCGTGCAAGAGAGATCGAGGTCCGCGACACCAAGAGAGCC				300
DB	241 ACCGTGGCCACCCCTGTACTGCGTGCAAGAGAGATCGAGGTCCGCGACACCAAGAGAGCC				300
QY	301 CTGAGCAAGATCGAGAGGAGCAGAACAAGTCCAGCAGAAAGATCCAGCAGGCCGAGGCC				360

DB 301 CTGACAAAGATCGAGAGAGAGAGAAAGTCCAGAGAGATCCAGAGCCGAGGCC 360  
QY 361 GCCGACAAAGGCAAGGTGAGCCAGAACTACCCCATGTCAGAACTCTGAGGCGCATG 420  
DB 361 GCCGACAAAGGCAAGGTGAGCCAGAACTACCCCATGTCAGAACTCTGAGGCGCATG 420  
QY 421 GTGCAACAGGCAATCAAGCCCCGCAACCTGAAAGGCTGGTGAAGGTGATCGAGAGAG 480  
DB 421 GTGCAACAGGCAATCAAGCCCCGCAACCTGAAAGGCTGGTGAAGGTGATCGAGAGAG 480  
QY 481 GCCTTCAGCCCCGAGGTGATCCCATGTTCAACGCGCTTGAGGCGAGGCGCCACCCCCAG 540  
DB 481 GCCTTCAGCCCCGAGGTGATCCCATGTTCAACGCGCTTGAGGCGAGGCGCCACCCCCAG 540  
QY 541 GACCTGAACAGATGTTGAACACCGTGGCGGCAACAGGCGCCATGCAAGTCTGAAG 600  
DB 541 GACCTGAACAGATGTTGAACACCGTGGCGGCAACAGGCGCCATGCAAGTCTGAAG 600  
QY 601 GACACCATCAACGAGAGAGGCGCGGAGTGGAGACCGCGTGCACCCCGTGCACCGCGGCCCC 660  
DB 601 GACACCATCAACGAGAGAGGCGCGGAGTGGAGACCGCGTGCACCCCGTGCACCGCGGCCCC 660  
QY 661 ATCGCCCCCGGCGAGATGCGCGAGCCCCGCGGCAAGCATCGCCGCAACCAAGCACACC 720  
DB 661 ATCGCCCCCGGCGAGATGCGCGAGCCCCGCGGCAAGCATCGCCGCAACCAAGCACACC 720  
QY 721 CTGCAAGAGAGATCGCTGATGACCAAGCAACCCCATCCCGTGGCGGCAATCTAC 780  
DB 721 CTGCAAGAGAGATCGCTGATGACCAAGCAACCCCATCCCGTGGCGGCAATCTAC 780  
QY 781 AAGCGGTGATCATCTGCGGCTGAAACAAAGATGTCGGATGTAACAGCCCGTGAACATC 840  
DB 781 AAGCGGTGATCATCTGCGGCTGAAACAAAGATGTCGGATGTAACAGCCCGTGAACATC 840  
QY 841 CTGGAATCAAGAGAGGCGCCCAAGAGAGCCCTTCCGAGTACGTGACCGCTTCTTCAAG 900  
DB 841 CTGGAATCAAGAGAGGCGCCCAAGAGAGCCCTTCCGAGTACGTGACCGCTTCTTCAAG 900  
QY 901 ACCCTGCGCGCGAGAGAGAGCAACCAAGAGAGGTTGAAGAACTGATGACCGACACCTCTG 960  
DB 901 ACCCTGCGCGCGAGAGAGAGCAACCAAGAGAGGTTGAAGAACTGATGACCGACACCTCTG 960  
QY 961 GTGCAAGAGCGCAACCCCGAGCTGCAAGAACCATCTGCGCGCTCTGCGCCCGCGGCGCAGC 1020  
DB 961 GTGCAAGAGCGCAACCCCGAGCTGCAAGAACCATCTGCGCGCTCTGCGCCCGCGGCGCAGC 1020  
QY 1021 CTGGAAGAGATGATGACCGCTGCGCAAGAGGCGTGGGCGGCGCCCAAGGCGCGCTG 1080  
DB 1021 CTGGAAGAGATGATGACCGCTGCGCAAGAGGCGTGGGCGGCGCCCAAGGCGCGCTG 1080  
QY 1081 CTGGCCGAGGCGATGAGCCAGGCAACCAACAGCGTATGATGACAGAAAGCAACTTCAAG 1140  
DB 1081 CTGGCCGAGGCGATGAGCCAGGCAACCAACAGCGTATGATGACAGAAAGCAACTTCAAG 1140  
QY 1141 GGGCCCCGCGCATGCTCAAGTGTCTCACTGCGGCAAGAGAGGCCCATGCGCCGCAAC 1200  
DB 1141 GGGCCCCGCGCATGCTCAAGTGTCTCACTGCGGCAAGAGAGGCCCATGCGCCGCAAC 1200  
QY 1201 TGCCGCGCCCCCGCAAGAGGCTGCTGAAAGTGGCGCAAGAGGGGCCACCAAGTGAAG 1260  
DB 1201 TGCCGCGCCCCCGCAAGAGGCTGCTGAAAGTGGCGCAAGAGGGGCCACCAAGTGAAG 1260  
QY 1261 GACTGCAACCGAGGCGCAGGCCCACTTCTGAGCAAGATCTGGGCCAGGCCCAAGGCGCGC 1320  
DB 1261 GACTGCAACCGAGGCGCAGGCCCACTTCTGAGCAAGATCTGGGCCAGGCCCAAGGCGCGC 1320  
QY 1321 CCGGCAACTTCTCTGCAAGAGCGGCGCGAGCCCAAGGCGCGCGCGAGAGCTTCCGC 1380  
DB 1321 CCGGCAACTTCTCTGCAAGAGCGGCGCGAGCCCAAGGCGCGCGCGAGAGCTTCCGC 1380  
QY 1381 TTCAGAGAGACCAACCCCGCGCAAGAGCAAGAGCAAGAGCCGCGAGACCTTGACCAAGC 1440  
DB 1381 TTCAGAGAGACCAACCCCGCGCAAGAGCAAGAGCAAGAGCCGCGAGACCTTGACCAAGC 1440

DB 1381 TTCAGAGAGACCAACCCCGCGCAAGAGCAAGAGCAAGAGCCGCGAGACCTTGACCAAGC 1440  
QY 1441 CTGAAGAGCTGTTCCGCAACGACCCCTGAGCCAGTAA 1479  
DB 1441 CTGAAGAGCTGTTCCGCAACGACCCCTGAGCCAGTAA 1479  
RESULT 4  
AX455904  
LOCUS AX455904 1479 bp DNA linear PAT 06-JUL-2002  
DEFINITION Sequence 20 from Patent WO0204493.  
ACCESSION AX455904  
VERSION AX455904.1 GI:21714896  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCES 1 zur Megede, J., Barnette, S.W., Engelbrecht, S. and van Rensburg, B.  
AUTHORS polynucleotides encoding antigenic hiv type c polypeptides,  
TITLE Patent: WO 0204493-A 20 17-JAN-2002;  
JOURNAL CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)  
FEATURES  
source Location/Qualifiers  
1..1479  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="synthetic gag coding sequence of HIV strain AF110965"  
ORIGIN  
Query Match 98.9%; Score 1463; DB 6; Length 1479;  
Best Local Similarity 99.3%; Pred. No. 2.2e-149;  
Matches 1469; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
QY 1 ATGGGCGCCCGCGCAGCATCTGCGCGCGCGCAAGCTGAGCGCTGGAGCGCATCCGC 60  
DB 1 ATGGGCGCCCGCGCAGCATCTGCGCGCGCGCAAGCTGAGCGCTGGAGCGCATCCGC 60  
QY 61 CTGCGCCCGCGCGCAAGAGTGTACATGATGAAGCACTGTGTGGCGCAGCGCGAG 120  
DB 61 CTGCGCCCGCGCGCAAGAGTGTACATGATGAAGCACTGTGTGGCGCAGCGCGAG 120  
QY 121 CTGGAAGATTCGCCCTGAAACCCCGCGCTGCTGAGAGCAAGGCTGCAAGCATC 180  
DB 121 CTGGAAGATTCGCCCTGAAACCCCGCGCTGCTGAGAGCAAGGCTGCAAGCATC 180  
QY 181 ATCCGCAAGTGCACCCCGCTGAGAGCGGCAAGGAGCTGAAGAGCTGTCAAC 240  
DB 181 ATCCGCAAGTGCACCCCGCTGAGAGCGGCAAGGAGCTGAAGAGCTGTCAAC 240  
QY 241 ACCGTGACCACTGTACTGCTGACAGAGATGAGTCCGCGACCAAGAGAGGCC 300  
DB 241 ACCGTGACCACTGTACTGCTGACAGAGATGAGTCCGCGACCAAGAGAGGCC 300  
QY 301 CTGGAAGAATCGAGAGAGCAAGTTCAGTCCAGAGCGGCGAGGCC 360  
DB 301 CTGGAAGAATCGAGAGAGCAAGTTCAGTCCAGAGCGGCGAGGCC 360  
QY 361 GCCGACAAAGGCAAGGTGAGCCAGAACTACCCCATGTCAGAACTGCAAGGCGCAGATG 420  
DB 361 GCCGACAAAGGCAAGGTGAGCCAGAACTACCCCATGTCAGAACTGCAAGGCGCAGATG 420  
QY 421 GTGCAACAGGCGCATCAGCCCGCAACCTGAAAGCTGAGTGAAGTATCGAGAGAG 480  
DB 421 GTGCAACAGGCGCATCAGCCCGCAACCTGAAAGCTGAGTGAAGTATCGAGAGAG 480  
QY 481 GCCTTCAGCCCGAGGTGATCCCATGTTCAACGCGCTGAGCGAGGCGCCACCCCCAG 540  
DB 481 GCCTTCAGCCCGAGGTGATCCCATGTTCAACGCGCTGAGCGAGGCGCCACCCCCAG 540  
QY 541 GACCTGAACAGATGTTGAACACCGTGGCGGCAACAGGCGCCATGCAAGTCTGAAG 600  
DB 541 GACCTGAACAGATGTTGAACACCGTGGCGGCAACAGGCGCCATGCAAGTCTGAAG 600



```

Db      541 GACCTGAACCATGCTGAACACCGTGGCGGCGCACAGGCCGCATGCAATGCTGAAG 600
Qy      601 GACACCATCAACGAGAGAGCGCGAGTGGGACCGCTGCAACCCGTGACAGCGCGCC 660
Db      601 GACACCATCAACGAGAGAGCGCGAGTGGGACCGCTGCAACCCGTGACAGCGCGCC 660
Qy      661 ATGCGCGCGCGCGAGATGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
Db      661 ATGCGCGCGCGCGAGATGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
Qy      721 CTGAGAGAGAGATGCGCTGATGACCAAGACCCCGCATCCCGTGGCGCATCTAC 780
Db      721 CTGAGAGAGAGATGCGCTGATGACCAAGACCCCGCATCCCGTGGCGCATCTAC 780
Qy      781 AAGCGGTGATCATCTGGGCTGAACAAGATCGTGCGATGTACAGCCCGTGAAGCATC 840
Db      781 AAGCGGTGATCATCTGGGCTGAACAAGATCGTGCGATGTACAGCCCGTGAAGCATC 840
Qy      841 CTGACATCAAGCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
Db      841 CTGACATCAAGCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
Qy      901 ACCGTGCGCGCGCGAGCAGAGCAACCGAGGTGAAGAACTGATGACCGCATCTGCTG 960
Db      901 ACCGTGCGCGCGCGAGCAGAGCAACCGAGGTGAAGAACTGATGACCGCATCTGCTG 960
Qy      961 GTGAGAAAGCCCAACCCCGAGCTGCAAGACCATCTGCGCGCTCTGCGCGCGCGCG 1020
Db      961 GTGAGAAAGCCCAACCCCGAGCTGCAAGACCATCTGCGCGCTCTGCGCGCGCGCG 1020
Qy      1021 CTGAGAGAGATGATGACCGCTGCGAGGCGTGGCGCGCGCGCGCGCGCGCGCGCTG 1080
Db      1021 CTGAGAGAGATGATGACCGCTGCGAGGCGTGGCGCGCGCGCGCGCGCGCGCTG 1080
Qy      1081 CTGCGCGAGCGCGATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1140
Db      1081 CTGCGCGAGCGCGATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1140
Qy      1141 GCGCGCGCGCGCGATCTGCAAGTGTCTCAACTGCGCGCGCGCGCGCGCGCGCGCG 1200
Db      1141 GCGCGCGCGCGCGATCTGCAAGTGTCTCAACTGCGCGCGCGCGCGCGCGCGCGCG 1200
Qy      1201 TGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1260
Db      1201 TGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1260
Qy      1261 GACTGCAACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1320
Db      1261 GACTGCAACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1320
Qy      1321 CCGCGCGCACTTCTGCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
Db      1321 CCGCGCGCACTTCTGCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
Qy      1381 TTGAGAGAGACCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1440
Db      1381 TTGAGAGAGACCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1440
Qy      1441 CTGAAGAGCTGTCTGCGCAACGACCCCTGAGCGCACTAA 1479
Db      1441 CTGAAGAGCTGTCTGCGCAACGACCCCTGAGCGCACTAA 1479

```

```

RESULT 5
LOCUS   AY181195               1485 bp    DNA             linear     SYN 02-OCT-2003
DEFINITION   Synthetic construct HIV-1-derived gag protein (gag) gene, complete
            cds.
ACCESSION   AY181195
VERSION     AY181195.1   GI:37413992
KEYWORDS    synthetic construct
SOURCE

```

```

ORGANISM   synthetic construct
REFERENCE   1 (bases 1 to 1485)
AUTHORS    Gao,F., Li,Y., Decker,J.M., Peyerl,F.W., Bibollet-Ruche,F.,
            Rodenburg,C.M., Chen,Y., Shaw,D.R., Allen,S., Musonda,R.,
            Shaw,G.M., Zajac,A.J., Letvin,N. and Hahn,B.H.
TITLE      Codon usage optimization of HIV type 1 subtype C gag, pol, env, and
            nef genes: in vitro expression and immune responses in
            DNA-vaccinated mice
JOURNAL    AIDS Res. Hum. Retroviruses 19 (9), 817-823 (2003)
REFERENCE   2 (bases 1 to 1485)
AUTHORS    Gao,F., Li,Y., Decker,J.M., Bibollet-Ruche,F., Rodenburg,C.M.,
            Chen,Y., Shaw,D.R., Allen,S., Musonda,R., Shaw,G.M., Zajac,A.J. and
            Hahn,B.H.
TITLE      Direct Submission
JOURNAL    Submitted (16-NOV-2002) Medicine, Duke University Medical Center,
            112 RPIII, Lasalle Street, Durham, NC 27710, USA
FEATURES
SOURCE

```

```

1. .1485
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="codon usage optimization of HIV-1 962M651"
1. .1485
/gene="gag"
1. .1485
/gene="gag"
/codon_start=1
/transl_table=1
/product="gag protein"
/protein_id="AA065393.1"
/db_xref="GI:37413993"
/translation="MGARASILRGKLDKWEKIRLPGGKRYWIKHLVWASRLERF
ALNPGILBTSBCKQIMKOLQPALQGTBELRSUNYNTVATLYVHGVEDRTKEALD
RIEENKIQKIQKIQKQQAADKVSQNYPIVQNLQGMVHOKSPRTLNAMWVYIEE
KAPSPBVIPIMTFALSEGATPDINTMLNTVGGHQAAMQMLKDTINEBAEMDRLEPVH
AGPIAPGMREPRGSDIAGTSTLQEQIAMTSPNPIPVGDIYRWIILGLNKIYRMV
SPVSLIDIKQKPEPRDYVDRFKTLRAEQATQEVKMMTDTLVONANPDCKTLK
ALPGATLEEMTACQGVGSPSHKARVLAEAMQTNVNIILQKSNFKGNKRMVCFN
CGEIGHIARNCAAPRKKGCKWGCKEGHQMDCTEERQANFLIKIWPSHKRGPNPLQNR
PEPTAPPAESFRFBETTPAPKQESKREALTSLKSLFGSDPLSQ"

```

## ORIGIN

```

Query Match      89.8%; Score 1328.2; DB 12; Length 1485;
Best Local Similarity 94.3%; Pred. No. 8e-135;
Matches 1401; Conservative 0; Mismatches 78; Indels 6; Gaps 2;

Qy      1 ATGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
Db      1 ATGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
Qy      61 CTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
Db      61 CTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
Qy      121 CTGAGAGAGTTCGCGCTGAAACCCCGCGCTGCGAGACCGAGGCGCTGCAAGCAGATC 180
Db      121 CTGAGAGCGCTTCGCGCTGAAACCCCGCGCTGCGAGACCGAGGCGCTGCAAGCAGATC 180
Qy      181 ATCCGCGAGCTGCAACCCCGCGCTGCGAGACCGCGCGCGCGCGCGCGCGCGCGCGCG 240
Db      181 ATGAAGCAGCTGCAACCCCGCGCTGCGAGACCGCGCGCGCGCGCGCGCGCGCGCGCG 240
Qy      241 ACCGTGCGCAACCTGTACTGCTGCGACGAGAAAGATCGAGTCCCGCGACCAAGAGAGCGC 300
Db      241 ACCGTGCGCAACCTGTACTGCTGCGACGAGGCGCTGAGAGTCCCGCGACCAAGAGAGCGC 300
Qy      301 CTGACACAGATCGAGAGAGACGAAACAAGTCCAGCAGAGATCCAGCAG---GCCGAG 357
Db      301 CTGACCCGATCGAGAGAGACGAAACAAGATCCAGCAGAGATCCAGCAGAAAAACCCAG 360
Qy      358 GCCGCGCAAGGCGCAGGTGAGCGCAAGACTAACCCCATGCTGCGAGAACTGCGAGGGCGCAG 417

```

```
Db 361 CAGGCCGCGGCGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACTGCGAGGGCCAG 420
Qy 418 ATGTGACACCAAGCCATCAGCCCCCGCACCCCTGAAACGCTGGGTGAAGGTGATCGAGAG 477
Db 421 ATGTGACACCAAGAGCTGAGCCCCCGCACCCCTGAAACGCTGGGTGAAGGTGATCGAGAG 480
Qy 478 AAGGCTTCAAGCCCCCGAAGGTGATCCCAATGTTCAACCGCCCTGAGCGAGGGCCACCCCC 537
Db 481 AAGGCTTCAAGCCCCCGAAGGTGATCCCAATGTTCAACCGCCCTGAGCGAGGGCCACCCCC 540
Qy 538 CAGGACCTGAAACAGATGTTGAAACACCGTGGCGGGCCACCGAGGCCCATGAGATGCTG 597
Db 541 CAGGACCTGAAACAGATGTTGAAACACCGTGGCGGGCCACCGAGGCCCATGAGATGCTG 600
Qy 598 AAGGACACCATCAACGAGGAGGCGCGCGAGTGGAGACCGCGTGCACCCCGTGCACGCGGC 657
Db 601 AAGGACACCATCAACGAGGAGGCGCGCGAGTGGAGACCGCGTGCACCCCGTGCACGCGGC 660
Qy 658 CCCATCGCCCCCGCGCAGATGCGCGAGGCCCCCGCGCGCAGACATCGCCGCGCACCAAGC 717
Db 661 CCCATCGCCCCCGCGCAGATGCGCGAGGCCCCCGCGCGCAGACATCGCCGCGCACCAAGC 720
Qy 718 ACCCTGCAAGAGCAAGATGCGCTGGATGACCAAGAACCCCCCAATCCCGTGGCGGACATC 777
Db 721 ACCCTGCAAGAGCAAGATGCGCTGGATGACCAAGAACCCCCCAATCCCGTGGCGGACATC 780
Qy 778 TACAAGCGGTGATCATCTGCGGCTGAAACAGATCGTGGGATGTACAGCCCCCGTGAGC 837
Db 781 TACAAGCGGTGATCATCTGCGGCTGAAACAGATCGTGGGATGTACAGCCCCCGTGAGC 840
Qy 838 ATCTGTGACATCAAGCAGGCGGCCCAAGGAGCCCTTCGCGACTACGTGAGACCGCTTCTTC 897
Db 841 ATCTGTGACATCAAGCAGGCGGCCCAAGGAGCCCTTCGCGACTACGTGAGACCGCTTCTTC 900
Qy 898 AAGACCTGCGCGCGCGAGCAGAGCAACCCAGAGGTGAAAGAACTGATGACCGACACCTTG 957
Db 901 AAGACCTGCGCGCGCGAGCAGAGCAACCCAGAGGTGAAAGAACTGATGACCGACACCTTG 960
Qy 958 CTGTGCAAGAACGCCAACCCCGCATGCAAGACCATCTGCGGCTCTGCGGCCCGCGCGCC 1017
Db 961 CTGTGCAAGAACGCCAACCCCGCATGCAAGACCATCTGCGGCTCTGCGGCCCGCGCGCC 1020
Qy 1018 AGCTGTGAGAGATGATGACCGCCTGCGCAGGCGGTGGCGGCCCGACCAAGGCCCGCC 1077
Db 1021 ACCCTGAGAGATGATGACCGCCTGCGCAGGCGGTGGCGGCCCGACCAAGGCCCGCC 1080
Qy 1078 GTGCTGCGCGAGGCGGATGAGCCAGGCCAACAC--AGCGTATGATGTCAGAGAGCAAC 1134
Db 1081 GTGCTGCGCGAGGCGGATGAGCCAGGCCAGCGTGAACATCTGATGCAAGAGAGCAAC 1140
Qy 1135 TTCAGAGGCCCCCGCGCATGCTCAAGTCTTCAACTGCGGCAAGAGAGGCCCATGCGCC 1194
Db 1141 TTCAGAGGCAACAAGCGCATGCTGAAGTCTTCAACTGCGGCAAGAGAGGCCCATGCGCC 1200
Qy 1195 CGCAACTGCGCGCCCCCGCGAAGAGGGCTGTGAAAGTGGCGCAAGAGAGGCCCAACAG 1254
Db 1201 CGCAACTGCGCGCCCCCGCGAAGAGGGCTGTGAAAGTGGCGCAAGAGAGGCCCAACAG 1260
Qy 1255 ATGAAGACTGCAACGAGCGCCAGGCCAACTTCTGCGCAAGATCTGGCCAGCCACAAG 1314
Db 1261 ATGAAGACTGCAACGAGCGCCAGGCCAACTTCTGCGCAAGATCTGGCCAGCCACAAG 1320
Qy 1315 GGGCGCGCGCGCAACTTCTGCAAGAGCGCCCGAGGCCCAACCGCCCCCGCGAGAGC 1374
Db 1321 GGGCGCGCGCGCAACTTCTGCAAGAGCGCCCGAGGCCCAACCGCCCCCGCGAGAGC 1380
Qy 1375 TTCCGCTTGAAGAGACCAACCGCCCGCGCAGAGCAGAGAGAGAGAGACCGCGAGACCTG 1434
Db 1381 TTCCGCTTGAAGAGACCAACCGCCCGCGCCCAAGAGAGAGAGAGAGACCGCGAGACCTG 1440
Qy 1435 ACCAGCTGAAGAGCTGTTCGCAACGAGACCCCTGAGCCAGTAA 1479
Db 1441 ACCAGCTGAAGAGCTGTTCGCAACGAGACCCCTGAGCCAGTAA 1485
```

```
RESULT 6
AX455983
LOCUS
DEFINITION
Sequence 99 from Patent WO0204493.
ACCESSION
AX455983
VERSION
AX455983.1
GI:21714967
KEYWORDS
SOURCE
synthetic construct
ORGANISM
synthetic construct
artificial sequences.
REFERENCE
1 zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.
AUTHORS
Polynucleotides encoding antigenic hiv type c polypeptides,
TITLE
Polypeptides and uses thereof
JOURNAL
Patent: WO 0204493-A 99 17-JAN-2002;
CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)
FEATURES
Location/Qualifiers
source
1..1491
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Gag_TV2_C_ZAopt"
ORIGIN
Query Match 87.1%; Score 1288.8; DB 6; Length 1491;
Best Local Similarity 93.0%; Pred. No. 1.5e-130;
Matches 1389; Conservative 0; Mismatches 87; Indels 18; Gaps 3;
Qy 1 ATGGGCGCGCGCGCGCAGCATCTGCGCGCGCGCGCAAGCTGAGCGCTGGAGCGCATCCGC 60
Db 1 ATGGGCGCGCGCGCGCAGCATCTGCGCGCGCGCGCAAGCTGAGCGCTGGAGCGCATCCGC 60
Qy 61 CTGCGCGCGCGCGCGCAAGAGTGTCTACATGATGAGGACCTGTGTGTGGGCCAGCGCGAG 120
Db 61 CTGCGCGCGCGCGCGCGCAAGAGTGTCTACATGATGAGGACCTGTGTGTGGGCCAGCGCGAG 120
Qy 121 CTGAGAGAGTTGCGCCCTGTAACCCCGGCGCTGTGAGAGACCGAGCGGCTGCAAGCATC 180
Db 121 CTGAGAGCGTTGCGCGGTGTAACCCCGGCGCTGTGAGAGACCGAGCGGCTGCGCCAGATC 180
Qy 181 ATCCGCAAGCTGCAACCCCGGCGCTGCAAGACCGGAGAGAGCTGAAGAGCGCTGTTAAC 240
Db 181 ATCAAGCAGCTGCAAGCCCGGCGCTGCAAGACCGGAGAGATCCGAGCGCTGTTAAC 240
Qy 241 ACCGTGGCCACCTGTACTGCGGTGCAAGAGATCGAGGTTCCGCAACCAAGAGGCC 300
Db 241 ACCGTGGCCACCTGTACTGCGGTGCAAGAGATCGAGGTTCCGCAACCAAGAGGCC 300
Qy 301 CTGACAAGATCGAGAGAGAGAGCAAGCAAGTGCACAGCAAGATCCAGAGGCCGAGGCC 360
Db 301 CTGACAAGATCGAGAGAGAGAGCAAGCAAGTGCACAGCAAGATCCAGAGGCCGAGGCC 360
Qy 361 GCCGACAAGGCGAAGGTGAGCCAGAACTACTCCATCTGTGCAAGACCTGCAAGGCCAGATG 420
Db 361 GCCGACAAG--AAGTGAAGCCAGAACTACTCCATCTGTGCAAGACCTGCAAGGCCAGATG 417
Qy 421 GTGACACGAGCCATCAGCCCCCGCACCCCTGAAACGCTTGGGTGAAGGTGATCGAGAGAG 480
Db 418 GTGACACGAGCCATCAGCCCCCGCACCCCTGAAACGCTTGGGTGAAGGTGATCGAGAGAG 477
Qy 481 GCCTTCAAGCCCGAGGTGATCCCATGTTCAACCGCCCTGAGCGAGGGCGCCACCCCCAG 540
Db 478 GCCTTCAAGCCCGAGGTGATCCCATGTTCAACCGCCCTGAGCGAGGGCGCCACCCCCAG 537
Qy 541 GACCTGAACAGATGTTGAACACCGTGGGCGGCCACAGGCGGCCATGCAAGTCTGAAG 600
Db 538 GACCTGAACAGATGTTGAACACCGTGGGCGGCCACAGGCGGCCATGCAAGTCTGAAG 597
Qy 601 GACACCATCAACGAGGAGCGCGCGAGTGGACCGCGTGCACCCCGTGCAACCGCGGCGCC 660
Db 598 GACACCATCAACGAGGAGCGCGCGAGTGGACCGCGTGCACCCCGTGCAACCGCGGCGCC 657
```

OY	661	ATGCCCCCCCCGATGCGCGAGCCCCCGGCGAGCGACATCGCTGGCAACCAAGCAACC	720
DB	658	GTGGCCCCCGGCGAGATGCGCGAGCCCCCGGCGAGCGACATCGCTGGCAACCAAGCAACC	717
OY	721	CTGCAGAGCAGATCGCTTGATGACCAAGCAACCCCCCATCCCCGTGGCGCATCTAC	780
DB	718	CTGCAGAGCAGATCGCTTGATGACCAAGCAACCCCCCATCCCCGTGGCGCATCTAC	777
OY	781	AGCGGTGATCATCTCGGCGCTGAACAAGATCGTCGAGTGTACAGCCCCGTGAGCATC	840
DB	778	AGCGGTGATCATCTCGGCGCTGAACAAGATCGTCGAGTGTACAGCCCCGTGAGCATC	837
OY	841	CTGCACATCAAGCAGGCGCCCAAGAGCCCCCTTCGCGACTACGTGGAACCGCTTCTCAAG	900
DB	838	CTGCACATCAAGCAGGCGCCCAAGAGCCCCCTTCGCGACTACGTGGAACCGCTTCTCAAG	897
OY	901	ACCTGCGCGCGCAGACAGCAACCAAGAGGTGAAGAACTGATGACCGACACCTGCTG	960
DB	898	ACCTGCGCGCGCAGACAGCAACCAAGAGGTGAAGAACTGATGACCGACACCTGCTG	957
OY	961	GTGCAGAACGCAACCCCGACTGCAAGACCATCTGCGCGCTCTCGCCCCGCGCGACG	1020
DB	958	GTGCAGAACGCAACCCCGACTGCAAGACCATCTGCGCGCTCTCGCCCCGCGCGACG	1017
OY	1021	CTGCAGAGATGATGACCGCTGCGCAGGCGTGGCGCGGCCCAAGCCAGCGCGGTG	1080
DB	1018	CTGCAGAGATGATGACCGCTGCGCAGGCGTGGCGCGGCCCAAGCGCGCGGTG	1077
OY	1081	CTGGCCGAGGCGATGAGCCAGGC---CAACACCAAGCTGATGATCAGAAGCAACTTC	1137
DB	1078	CTGGCCGAGGCGCATGAGCCAGGCCAACACACCAAGCTGATGATCAGAAGCAACTTC	1137
OY	1138	AAAGGCCCCCGCGCATCGTCAAGTCTTCAACTGCGGCAAGAGGGCCACATGCCCCG	1197
DB	1138	AAAGGCCCCCGCGCGCGTGAAGTCTTCAACTGCGGCGGCGAGGGGCCACATGCCCCG	1197
OY	1198	AACTGCGCGCGCCCCCGCAAGAGGGCTGTGAAGTGCAGAGGCGCAAGAGGCCCAAGATG	1257
DB	1198	AACTGCGCGCGCCCCCGCAAGCGGGCTGTGAAGTGCAGAGGCGCAAGAGGCCCAAGATG	1257
OY	1258	AAAGACTGCACCGAGCGCCCAAGCCAACTTCCTGGGCAAGATCTGGCCCAAGGGG	1317
DB	1258	AAAGACTGCACCGAGCGCCCAAGCCAACTTCCTGGGCAAGATCTGGCCCAAGGGG	1317
OY	1318	CGCCCCGCGCACTTCTCTGCAAGATCGCCCCGAGCCCCAACCGCCCCCGCGAGACTTC	1377
DB	1318	CGCCCCGCGCACTTCTCTGCAAGATCGCCCCGAGCCCCAACCGCCCCCGCGAGCCACC	1377
OY	1378	CGCTTCGAGG-----AGACCAACCCCGGCCAAGAGCAAGAGCAAGAACCGC	1425
DB	1378	CGCTTCGAGG-----AGACCAACCCCGGCCAAGAGCAAGAGCAAGAACCGC	1437
OY	1426	GAGAACCTGACCAAGCCTGAAGAGCTGTTGCGCAACGACCCCTGAGCCAGTAA	1479
DB	1438	GAGCCCTTGACCAAGCCTGAAGAGCTGTTGCGCAACGACCCCTGAGCCAGTAA	1491

RESULT 7			
AX455935	AX455935	1494 bp	DNA linear PAT 06-JUL-2002
LOCUS	Sequence S1 from Patent WO0204493.		
DEFINITION	AX455935		
ACCESSION	AX455935		
VERSION	AX455935.1 GI:21714919		
KEYWORDS	.		
SOURCE	synthetic construct		
ORGANISM	synthetic construct		
	artificial sequences.		

REFERENCE  
1  
AUTHORS zur Megede, J., Barnett, S.W., Engelbrecht, S. and van Rensburg, B.  
TITLES Polynucleotides encoding antigenic hiv type c polypeptides,  
polypeptides and uses thereof  
JOURNAL Patent: WO 0204493-A 51 17-JAN-2002;

```

FEATURES
    source
        1. 1494
            /organism="synthetic construct"
            /mol_type="unassigned DNA"
            /db_xref="taxon:32630"
            /note="HIV Type C Gag optimized"
ORIGIN

```

Query Match	87.1%	Score 1288;	DB 6;	Length 1494;
Best Local Similarity	92.9%	Pred. No. 1.0e-130;		
Matches 1393; Conservative	0;	Mismatches 80;	Indels 27;	Gaps 3;

QY	1	ATGGGCGCCCGGCAGCATCTCTGCGCGCGGCAAGCTGACGCGCTGGGAGCGCATCCGC	60
Db	1	ATGGGCGCCCGGCAGCATCTCTGAGCGCGGCAAGCTGACGCGCTGGGAGCGCATCCGC	60
QY	61	CTGCGCCCCCGCGCAAGAGTGTACATGATGAAGCACTGTGTGGGCCAGCCGCGAG	120
Db	61	CTGCGCCCCCGCGCAAGAGCACTACATGCTGAAGCACTGTGTGTGGGCCAGCCGCGAG	120
QY	121	CTGAGAAGATTCCGCTCTGAACCCCGGCTGTGGAGACCAGCGAGGCTGCAAGCAGATTC	180
Db	121	CTGAGAGCGCTTCGCTCTGAACCCCGGCTGTGTGGAGACCAGCGAGGCTGCAAGCAGATTC	180
QY	181	ATCCGCCAGCTGCACCCCGCCTTGCAGACCGGCAAGGAGCTGAAGAGCTGTTCAAC	240
Db	181	ATCAAGCAGCTGCAGCCCGCCTTGCAGACCGGCAAGGAGCTGCGCAGCTGTTCAAC	240
QY	241	ACCGTGGCCACCTCTGTACTGCGTGCACGAGAGATTCAGGTCCGCGCACCAAGAGGCC	300
Db	241	ACCGTGGCCACCTCTGTACTGCGTGCACAGGGGCATTCAGGTCCGCGCACCAAGAGGCC	300
QY	301	CTGGAACAAGATCGAGGAGGAGGACAGAACAAATGTCAGCAGAGATCCAGCAGGCCGAGGCC	360
Db	301	CTGGAACAAGATCGAGGAGGAGGACAGAACAAATGTCAGCAGAGATCCAGCAGGCCGAGGCC	360
QY	361	GCCGCAAGGGCAAGGTGAGCCCAACTACCCCATCTGTGCAGAACCTGCAGGGCCAGATG	420
Db	361	GCCGCAAGG--AAGGTGAGCCAGAACTACCCCATCTGTGCAGAACCGCCAGGGCCAGATG	417
QY	421	GTGCACCAAGGCATCAGCCCCCGCACCTTGAACGCGCTGGTGAAGGTGATCGAGAGAAAG	480
Db	418	GTGCACCAAGGCATCAGCCCCCGCACCTTGAACGCGCTGGATCAAGGTGATCGAGAGAAAG	477
QY	481	GCCCTCAGCCCGAGGTGATCCCATGTTCAACCGCCTGAGCGAGGCGCCACCCCCCAG	540
Db	478	GCCCTCAACCCCGAGAGATCCCATGTTCAACCGCCTGAGCGAGGCGCCACCCCCCAG	537
QY	541	GACCTGAACACGATGTTGAACACCGTGGCGGCCAACAGGCCGCAATGCAGATGCTGAAG	600
Db	538	GACCTGAACACCATGCTGGAACACCTGGCGGCCAACAGGCCGCAATGCAGATGCTGAAG	597
QY	601	GACACCATCAACGAGAGGCGCGCGAGTGGAACCGCGTGCACCCCGTGCACGCGCGCCC	660
Db	598	GACACCATCAACGAGAGGCGCGCGAGTGGAACCGCACCCACCCCGTGCACGCGCGCCC	657
QY	661	ATCGCCCCCGGCAGATGCGCGAGGCCCCGCGCGCAGCATCGCCGCAACCAAGCAGACC	720
Db	658	GTGGCCCCCGGCAGATGCGCGAGGCCCCGCGCGCAGCATCGCCGCAACCAAGCAGACC	717
QY	721	CTGCAGAGAGCATCGCCTGATGAACAGCAACCCCCCATCCCGGTGGGGCAGCATCTAC	780
Db	718	CTGCAGAGAGCATCGCCTGATGAACAGCAACCCCCCATCCCGGTGAAGCAGCATCTAC	777
QY	781	AAGCGGTGATCATCTGGGCTGAACAAGATCGTGGGATGTACAGCCCCGTGAGCATC	840
Db	778	AAGCGGTGATCATCTGGGCTGAACAAGATCGTGGGATGTACAGCCCCGTGAGCATC	837
QY	841	CTGCACATCAACGAGGGCCCAAGAGGCCCTTCCGCGACTACGTGACCGCTTCTTCAAG	900
Db	838	CTGCACATCAACGAGGGCCCAAGAGGCCCTTCCGCGACTACGTGACCGCTTCTTCAAG	897

OY		901	ACCTGCGCGCCGAGCAGAGCAACCAGAGGTGAAGAACTGGATGACCCGACTGTCG	960
Db		898	ACCTGCGCGCCGAGCAGAGGCCAACCCAGGACGTGAAGAACTGGATGACCCGACTGTCG	957
OY		961	GTCGAGAATCGCCAACCCCGCATGTCGAAGACCATCTGCGCGCTCTGCGCCCCGCGCCAGC	1020
Db		958	GTCGAGAATCGCCAACCCCGCATGTCGAAGACCATCTGCGCGCTCTGCGCCCCGCGCCAGC	1017
OY		1021	CTGAGAGAGATGATGACCGCGCTGCGCAGGCGCTGGGCGCGGCCCAAGCAAAGGCCGCGTG	1080
Db		1018	CTGAGAGAGATGATGACCGCGCTGCGCAGGCGCTGGGCGCGGCCCAAGCAAAGGCCGCGTG	1077
OY		1081	CTGGCCGAGGCGATGAGCCGAGGCCAACAACCGAGTGTATGATGACGAAGACCAATTCAAG	1140
Db		1078	CTGGCCGAGGCGATGAGCCGAGGCCAACAACCGAGTGTATGATGACGAAGACCAATTCAAG	1137
OY		1141	GSCCCCOCGCGCATGCTCAAGTGCTTCAACTGCGGCAGAGAGGSCCAATGSCCCGCAAC	1200
Db		1138	GSCCAGCAACCGCATCATCAAGTGCTTCAACTGCGGCAGAGAGGSCCAATGSCCCGCAAC	1197
OY		1201	TGCGCGCGCCCCCGGCAGAGAGGGCTGCTGGAAGTGCGGCGAAGAGGGCCACAGATGAAG	1260
Db		1198	TGCGCGCGCCCCCGGCAGAGAGGGCTGCTGGAAGTGCGGCGAAGAGGGCCACAGATGAAG	1257
OY		1261	GACTGCACACCGAGCGCCAGGSCCAATTCTTGCGCAAGATCTGSCCCAGCCACAAGGGCCGC	1320
Db		1258	GACTGCACACCGAGCGCCAGGSCCAATTCTTGCGCAAGATCTGSCCCAGCCACAAGGGCCGC	1317
OY		1321	CCCGGCAACTTCTCTGCAAGCGCGCC-----CCGAGCCCAACCGCC	1359
Db		1318	CCCGGCAACTTCTCTGCAAGCGCGCCCGAGCCCAACCGCCCCCGCGAGCCCAACCGCC	1377
OY		1360	CCCCCGCGCGAGAGCTTCCGCTTTCGAGAGACCAACCCCGCGCAGAGAGAGAGCAAG	1419
Db		1378	CCCCCGCGCGAGAGCTTCCGCTTTCGAGAGACCAACCCCGTGGCCCGCAAGAG--AAG	1434
OY		1420	GACCGCGAGACCTGACCAAGCCTGAAGAGCCTGTTCCGCAACGACCCCTGAGCCAGTA	1479
Db		1435	GACCGCGAGACCTGACCAAGCCTGAAGAGCCTGTTCCGCAACGACCCCTGAGCCAGTA	1494
<b>RESULT 8</b>				
AX455888				
LOCUS		AX455888	1509 bp	DNA linear PAT 06-JUL-2002
DEFINITION		Sequence 4 from Patent WO0204493.		
ACCESSION		AX455888		
VERSION		AX455888.1 GI:21714881		
KEYWORDS		synthetic construct		
SOURCE		artificial sequences.		
ORGANISM		synthetic construct		
REFERENCE		1 zur Megede,U., Barnett,S.W., Engelbrecht,S. and van Rensburg,B.		
AUTHORS		Polynucleotides encoding antigenic hiv type c polypeptides,		
TITLE		Patent: WO 0204493-A 4 17-JAN-2002;		
JOURNAL		CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)		
FEATURES		Location/Qualifiers		
source		1..1509		
		/organism="synthetic construct"		
		/mol_type="unassigned DNA"		
		/db_xref="taxon:32630"		
		/note="Synthetic Gag of HIV strain AF110967"		

Query Match	86.3%	Score 1276.8;	DB 6;	Length 1509;
Best Local Similarity	92.5%	Pred. No. 2.9e-129;		
Matches 1399;	Conservative	0;	Mismatches 77;	Indels 36;
				Gaps 4;

[illegible]

OY	61	CTGCGCCCCCGGCAAGAAAGTGTCTACATGATGAAAGCACTGTGTGTGGGCCAGCCGGGAG	120
Db	61	CTGCGCCCCCGGCGGCAAGAAAGCACTACATGTCTGAAGCACTGTGTGTGGGCCAGCCGGGAG	120
OY	121	CTGGAGAAAGTTTCGGCCCTGTAACCCCGGGCTGTGTGGAGACCAGCGAGGGCTGTCAAGCAGATC	180
Db	121	CTGGAGAGGGCTTTCGGCCCTGTAACTCCGGGCTGTGTGGAGACCGCGAGGGCTGTCAAGCAGATC	180
OY	181	ATCCGCCAGCTGTCAACCCCGCCCTGCAGACCGGCAGCGAGAGACTGAAGAGCTGTTCAC	240
Db	181	ATGAAGCAGCTGTCAAGCCCGCCCTGCAGACCGGCAGCGAGAGCTGTGCAGAGCTGTACAAAC	240
OY	241	ACCGTGGCCACCCCTGTACTGTGTGTCAACGAGAAAGATCGAGGTTCGGCAGACCAAGAGGGCC	300
Db	241	ACCGTGGCCACCCCTGTACTGTGTGTCAACGAGCCGGCATCGAGGTTCGGCAGACCAAGAGGGCC	300
OY	301	CTGGACAAGATCTGAAGAGGAGCAGAACAGATGTCCAGCAGAAAGATCCAGCAGGCCGAGGCC	360
Db	301	CTGGACAAGATCTGAAGAGGAGCAGAACAGATGTCCAGCAGAAAGATCCAGCAGGCCGAGGCC	360
OY	361	GCCGACAAAGGGGCAAGGTGAGCCAGAACTACCCCATCTGTGTCAAGAACTGTCAAGGGCCAGATG	420
Db	361	GCCGAC--GGCAAGGTGAGCCAGAACTACCCCATCTGTGTCAAGAACTGTCAAGGGCCAGATG	417
OY	421	GTGCACCAAGGCCATCAGCCCCCGGCAACCCTGAACGCTGTGGGTGAAGGTGATTCAGAGAGAG	480
Db	418	GTGCACCAAGGCCATCAGCCCCCGGCAACCCTGAACGCTGTGGGTGAAGGTGATTCAGAGAGAG	477
OY	481	GCCCTCAGCCCCCGAGGTGATTCCTCATGTTCACCGCCCTGAGCGAGGGGCGCCACCCCCAG	540
Db	478	GCCCTCAGCCCCCGAGGTGATTCCTCATGTTCACCGCCCTGAGCGAGGGGCGCCACCCCCAG	537
OY	541	GACCTGAACACGATGTTGAACACCGTGGGGCGGCCACCAAGGCCCATGTGACATGCTGAAG	600
Db	538	GACCTGAACACGATGTTGAACACCGTGGGGCGGCCACCAAGGCCCATGTGACATGCTGAAG	597
OY	601	GACACCATCAACGAGAGGAGCGCGCCGAGTGGGACCGCGTGCACCCCGTGCACGCGCGGGCCC	660
Db	598	GACACCATCAACGAGAGGAGCGCGCCGAGTGGGACCGCGCTGCACCCCGTGCACGCGCGGGCCC	657
OY	661	ATGCCCCCCCCGCAAGATGCGCGAGCCCGCGCGCAGCGACATGCGCGGCAACCAAGCACC	720
Db	658	GTGGCCCCCGGCAAGATGCGCGAGCCCGCGCGCAGCGACATGCGCGCGGCAACCAAGCACC	717
OY	721	CTGCAGAGAGCAGATGCGCTGTGATGACCAAGCAACCCCCCATTCCTGTGGGCGCATCTAC	780
Db	718	CTGCAGAGAGCAGATGCGCTGTGATGACCAAGCAACCCCCCGTGTGGGCGCATCTAC	777
OY	781	AAGCGGTGATCATCTCGGGCTGAACAAGATCGTGCAGATGTACAGCCCCGTGAGCATC	840
Db	778	AAGCGGTGATCATCTCGGGCTGAACAAGATCGTGCAGATGTACAGCCCCGTGAGCATC	837
OY	841	CTGGACATCAAGCAGGGCCCCCAAGAGGCCCTTTCGGCGACTACGTGACCCGCTTCTCAAG	900
Db	838	CTGGACATCCGCGCAGGGCCCCCAAGAGGCCCTTTCGGCGACTACGTGACCCGCTTCTCAAG	897
OY	901	ACCCTGCGCGCGGAGCAGAGCAACCCAGAGGCTGAAGAACTGATGACCGAGACCTGTGCTG	960
Db	898	ACCCTGCGCGCGGAGCAGAGCAACCCAGAGGCTGAAGAACTGATGACCGAGACCTGTGCTG	957
OY	961	GTGCAGAACGCAACCCCGACTGTCAAGACCATCTGCGCGCTCTCGGCCCCCGGCGCAAC	1020
Db	958	GTGCAGAACGCAACCCCGACTGTCAAGACCATCTGCGCGCTCTCGGCCCCCGGCGCAAC	1017
OY	1021	CTGGAGGAGATGATGACCGGCTGTCCAGGGGCTGTGGGGCGGCCAGCCAGAGGCCGCGTG	1080
Db	1018	CTGGAGGAGATGATGACCGGCTGTCCAGGGGCTGTGGGGCGGCCAGAGGCCGCGTG	1077
OY	1081	CTGGCCGAGGCGGATGAGCCAGGCCCAACACC--AGCGTGAATGATGCAGAAAGCAACTTC	1137
Db	1078	CTGGCCGAGGCGGATGAGCCAGGCCCAACAGCGTGAACATCATGATGCAGAAAGCAACTTC	1137
OY	1138	AAGGGCCCCCGGCGCATCTGTCAAGTGTCTTCACTGCGGCAAGAGAGGGCCACATCGCCCGC	1197

Db 1078 CTGGCCGAGGCGATGAGCCAGGCCCAACGCTGAACATCATGATGCAAGAAAGCAACTTC 1137

DY 1138 AAGGGCCCCCGGCGGCATTCGTCAAGTGTCTTCAACTGCGGCAAGAGAGGGCCACATCGCCGCG 1197



```
Db 1138 AAGGGCCCCCGCGCAACGTCAAGTCTTCACTGCGGCAAGAGGGCCACATCGCCAAG 1197
QY 1198 AACTGCGCGCCCCCGCGCAAGAGGCTGCTGGAAGTGCAGGCAAGAGGGCCACAGATG 1257
Db 1198 AACTGCGCGCCCCCGCGCAAGAGGCTGCTGGAAGTGCAGGCAAGAGGGCCACAGATG 1257
QY 1258 AAGACTGCAACCGAGGCGCAAGGCTTCTGCGCAAGATCTGCGCCAGCCACAAGGGC 1317
Db 1258 AAGACTGCAACCGAGGCGCAAGGCTTCTGCGCAAGATCTGCGCCAGCCACAAGGGC 1317
QY 1318 CGCCCCCGCACTTCTGCAAGAGCGG-----CGCCGAGCCCAACCGCC 1359
Db 1318 CGCCCCCGCACTTCTGCAAGAGCGG-----CGCCGAGCCCAACCGCC 1377
QY 1360 CCCCCCGCGGAGAGCTTCCGCTTGAGAGAACCAACCCCGCGCCAGAGCAAGAGCAAG 1419
Db 1378 CCCCCCGCGGAGAGCTTCCGCTTGAGAGAACCAACCCCGCGCCAGAGCAAGAGCAAG 1437
QY 1420 G-----ACCGCGAGAACCTTGACCAAGCTTGAAGAGCTTGTTCGGCAACGACCC 1467
Db 1438 GACCGCGAGCCCTACCGCGAGCCCTTGACCGCCCTTGCGGAGCTTGTTCGGCAAGCGCC 1497
QY 1468 CTGAGCCCACTAA 1479
Db 1498 CTGAGCCCACTAA 1509
```

```
RESULT 9
LOCUS AX468544 1509 bp DNA linear PAT 16-JUL-2002
DEFINITION Sequence 64 from Patent WO0226209.
ACCESSION AX468544
VERSION AX468544.1 GI:21901374
KEYWORDS
SOURCE Human immunodeficiency virus 1 (HIV-1)
ORGANISM Human immunodeficiency virus 1
          Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
          lentivirus group.
REFERENCE
  1 O'Hagan,D., Otten,G., Donnelly,J.J., Polo,J.M., Barnett,S.,
    Singh,M., Ulmer,J. and Dubensky,T.W.
    Microparticles for delivery of the heterologous nucleic acids
    Patent: WO 0226209-A 64 04-APR-2002;
    CHIRON CORPORATION (US)
FEATURES
  source
    1..1509
    /organism="Human immunodeficiency virus 1"
    /mol_type="unassigned DNA"
    /db_xref="taxon:11676"
ORIGIN
```

```
Query Match 86.2%; Score 1275.2; DB 6; Length 1509;
Best Local Similarity 92.5%; Pred. No. 4.3e-129;
Matches 1398; Conservative 0; Mismatches 78; Indels 36; Gaps 4;
```

```
QY 1 ATGGGCGCGCGCGCGCATCTGCGCGCGCGCAAGCTGAGCGCTGGAGCGCATCCGC 60
Db 1 ATGGGCGCGCGCGCGCGCATCTGCGCGCGCGCAAGCTGAGCGCTGGAGCGCATCCGC 60
QY 61 CTGCGCGCGCGCGCGGAGAGTGTCTAATGTAAGCACTGTGTGGCGCAGCGCGAG 120
Db 61 CTGCGCGCGCGCGCGGAGAGTGTCTAATGTAAGCACTGTGTGGCGCAGCGCGAG 120
QY 121 CTGGAAGAAGTTCCGCTGAACCCCGGCTGTGGAAGCAAGCGAGGCTGCAAGCATC 180
Db 121 CTGGAAGGCTTCCGCTGAACCCCGGCTGTGGAAGCAAGCGAGGCTGCAAGCATC 180
QY 181 ATCCCGCAGCTGCAACCCCGCTGTGAGACCGGCAAGAGAGCTGAAGAGCTTCAAC 240
Db 181 ATGAAGAGCTGCAAGCGCGCTGTGAGACCGGCAAGAGAGCTGCGAGCTGTACAAC 240
QY 241 ACCGTGCGCAACCTGTACTGCTGTCAACGAGAAGATCGAGTCCGCGACCAAGAGGCC 300
```

```
Db 241 ACCGTGCGCAACCTGTACTGCTGTCAACGCGGCGCATGAGTCCGCGACCAAGAGGCC 300
QY 301 CTGCAACAAGATCGAGAGAGCAGAACAGTCCAGCAGAGATCCAGCAGCGCGAGGCC 360
Db 301 CTGCAACAAGATCGAGAGAGCAGAACAGTCCAGCAGAGATCCAGCAGAGATCCAGCAG 360
QY 361 GCCGACAAGGCGAAGTGAAGCCAGAACTTACCCCATCTGTGAGAACTGTGAGGAGAG 420
Db 361 GCCGAC---GCGAAGGTGAGCCAGAACTTACCCCATCTGTGAGAACTGTGAGGAGAG 417
QY 421 GTGCACCAAGGCGCATCAGCCCGCGCAACCTGTAACGCTGGGTGAAGGTGATGAGAGAG 480
Db 418 GTGCACCAAGGCGCATCAGCCCGCGCAACCTGTAACGCTGGGTGAAGGTGATGAGAGAG 477
QY 481 GCCTTCAAGCCCGAGGTGATCCCATGTTCACCGCCCTGAGCGAGGCGCGCACCCCGCAG 540
Db 478 GCCTTCAAGCCCGAGGTGATCCCATGTTCACCGCCCTGAGCGAGGCGCGCACCCCGCAG 537
QY 541 GACCTGAACAAGATGTTGAACACCGTGGGCGGCCACAGGCGCCATGCAAGTCTGAAG 600
Db 538 GACCTGAACAAGATGTTGAACACCGTGGGCGGCCACAGGCGCCATGCAAGTCTGAAG 597
QY 601 GACACCATCAACGAGAGGCGCGCGAGTGGGACCGCGTGCAACCCCGTGACGCGCGGCC 660
Db 598 GACACCATCAACGAGAGGCGCGCGAGTGGGACCGCGTGCAACCCCGTGACGCGCGGCC 657
QY 661 ATGCCCCCGCGCGAGATGCGCGGAGGCCCGCGCGAGCGACATCGCGCGGACCAACAGCAC 720
Db 658 GTGCCCCCGCGCGAGATGCGCGGAGGCCCGCGCGAGCGACATCGCGCGGCGCACAGCAC 717
QY 721 CTGCAGAGAGAGATCGCTGTGATGACCAAGCAACCCCGCATCCCGGTGGCGCATCTAC 780
Db 718 CTGCAGAGAGAGATCGCTGTGATGACCAAGCAACCCCGCATCCCGGTGGCGCATCTAC 777
QY 781 AAGCGGTGATCATCTCGGCGCTGAACAAGATCGTGGGATGTACAGCCCGTGAGCATC 840
Db 778 AAGCGGTGATCATCTCGGCGCTGAACAAGATCGTGGGATGTACAGCCCGTGAGCATC 837
QY 841 CTGCACATCAACGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
Db 838 CTGCACATCGCGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 897
QY 901 ACCCTGCGCGCGCGAGCAGAGCAACCAAGAGGTGAAGAACTGTGATGACCGACACCTGCTG 960
Db 898 ACCCTGCGCGCGCGAGCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 957
QY 961 GTGCAGAAAGCGCAACCCCGCATGTCAAGACATCTGCGCGCTCTGCGCCCGCGCGCGAGC 1020
Db 958 GTGCAGAAAGCGCAACCCCGCATGTCAAGACATCTGCGCGCTCTGCGCCCGCGCGCGAGC 1017
QY 1021 CTGAGAGAGATGATGACCGCGCTGCGAGGCGGTGGCGCGCGCGCGCGCGCGCGCGGTG 1080
Db 1018 CTGAGAGAGATGATGACCGCGCTGCGAGGCGGTGGCGCGCGCGCGCGCGCGCGGTG 1077
QY 1081 CTGCGCGAGGCGATGAGCGCGCGCAACAC--AGCGTATGATGCAAGAGCAACTTC 1137
Db 1078 CTGCGCGAGGCGATGAGCGCGCGCAACAGCGTGAACATATGATGCAAGAGCAACTTC 1137
QY 1138 AAGGCGCGCGCGCGCATGCTCAAGTGTCTCAACTGCGCGCAAGAGAGGCGCACATCGCCGC 1197
Db 1138 AAGGCGCGCGCGCGCAAGCTCAAGTGTCTCAACTGCGCGCAAGAGAGGCGCACATCGCCAG 1197
QY 1198 AACTGCGCGCGCGCGCGCAAGAGGCTGTGGAAGTGCAGAGGCGCGCACAGATG 1257
Db 1198 AACTGCGCGCGCGCGCGCAAGAGGCTGTGGAAGTGCAGAGGCGCGCACAGATG 1257
QY 1258 AAGACTGCAACCGAGGCGCAAGGCTTCTGCGCAAGATCTGCGCCAGCCACAAGGGC 1317
Db 1258 AAGACTGCAACCGAGGCGCAAGGCTTCTGCGCAAGATCTGCGCCAGCCACAAGGGC 1317
QY 1318 CGCCCCCGCACTTCTGCAAGAGCGG-----CGCCGAGCCCAACCGCC 1359
```



Db 1318 CGCCCCGCACTTCTGAGAACCGCAGCAGCCCGCCGCCACCGTGCACCGCC 1377  
QY 1360 CCCCCCGGAGAGCTTCCGCTTGAGAGACCAACCCCGCCAGAGAGAGAG 1419  
Db 1378 CCCCCCGGAGAGCTTCCGCTTGAGAGACCAACCCCGCCAGAGAGAGAG 1437  
QY 1420 G-----ACCGGAGACCTGACCACTGAGAGAGCTGTTGGCAAGACCC 1467  
Db 1438 GACCGGAGACCTTACCGGAGACCTGACCTGCGCAGCTGTGCGAGCGGCC 1497  
QY 1468 CTGAGCCAGTAA 1479  
Db 1498 CTGAGCCAGTAA 1509

RESULT 10  
AX468548 1509 bp DNA linear PAT 16-JUL-2002  
LOCUS AX468548 Sequence 68 from Patent WO0226209.  
DEFINITION AX468548  
ACCESSION AX468548 GI:21901378  
VERSION AX468548.1  
KEYWORDS Human immunodeficiency virus 1 (HIV-1)  
SOURCE Human immunodeficiency virus 1  
ORGANISM Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate  
lentivirus group.

REFERENCE 1  
AUTHORS O'Hagan,D., Otten,G., Donnelly,J.J., Polo,J.M., Barnett,S.,  
Singh,M., Umer,J. and Dubensky,T.W.  
TITLE Microparticles for delivery of the heterologous nucleic acids  
JOURNAL Patent: WO 0226209-A 68 04-APR-2002;  
CHIRON CORPORATION (US)  
FEATURES Location/Qualifiers  
Source 1. 1509  
/organism="Human immunodeficiency virus 1"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:11676"

ORIGIN  
Query Match 86.1%; Score 1273.6; DB 6; Length 1509;  
Best Local Similarity 91.8%; Pred. No. 6.4e-129;  
Matches 1388; Conservative 12; Mismatches 76; Indels 36; Gaps 4;

QY 1 ATGGGCGCGCGCGCAGCATCTGCGCGCGCGCAGCTGACGCTGGAGCGCATCCGC 60  
Db 1 ATGGGCGCGCGCGCAGCATCTGCGCGCGCGCAGCTGACGCTGGAGAGATCCGC 60  
QY 61 CTGCGCGCGCGCGCAGAGAGTCTATCATGATGAAGCACTGTGTGGCGCGCGAG 120  
Db 61 CTGCGCGCGCGCGCAGAGAGCACTACATGCTGAAGCACTGTGTGGCGCGCGAG 120  
QY 121 CTGAGAGAGTTCGCTTGAACCCCGCGCTGTGAGAGACCAAGGCTGCAAGCATC 180  
Db 121 CTGAGAGGCTTCGCTTGAACCCCGCGCTGTGAGAGACCGCGGCTGCAAGCATC 180  
QY 181 ATCCGCGAGCTGCAACCCCGCGCTGCAAGCGCGAGAGCTGAAGAGCTTCAAC 240  
Db 181 ATGAGCAGCTGCAACCCCGCGCTGCAAGCGCGAGAGCTGCGCAGCTGTACAAC 240  
QY 241 ACCGTGCGCACTCTGTACTGCTGTCAGAGAGATGAGGTCCGCGCACAAGAGGCC 300  
Db 241 ACCGTGCGCACTCTGTACTGCTGTCAGAGCGCGGCACTGAGGTSCGCGACAAGAGGCC 300  
QY 301 CTGACCAAGATCGAGAGAGAGAGCAAGTGCAGAGAGATCCAGAGGCGGAGGCC 360  
Db 301 CTGACCAAGATCGAGAGAGAGAGCAAGTGCAGAGAGATCCAGAGGCGGAGAGGCC 360  
QY 361 GCCGCAAGGCGAGAGAGAGAGCAAGTGCAGAGAGATCCAGAGGCGGAGAGGCC 420  
Db 361 GCCGCAAGGCGAGAGAGAGAGCAAGTGCAGAGAGATCCAGAGGCGGAGAGGCC 420  
QY 421 GTGCAACGAGGCGATCAGCCCGCGCAGCTGAAAGCTTGGTGAAGGTGATGAGAGAG 480  
Db 421 GTGCAACGAGGCGATCAGCCCGCGCAGCTGAAAGCTTGGTGAAGGTGATGAGAGAG 480

Db 418 GTGCAACGAGGCGATCAGCCCGCGCAGCTGAAAGCTTGGTGAAGGTGATGAGAGAG 477  
QY 481 GCCTTACGCGCGCGAGGTGATCCCATGTTTCAACCGCTGAGAGAGGCGCCACCCCGAG 540  
Db 478 GCCTTACGCGCGCGAGGTGATCCCATGTTTCAACCGCTGAGAGAGGCGCCACCCCGAG 537  
QY 541 GACTGAACACGATGTTGAACACCGTGGGCGCCACAGGCGCCATGACATGCTGAAG 600  
Db 538 GACTGAACACGATGTTGAACACCGTGGGCGCCACAGGCGCCATGACATGCTGAAG 597  
QY 601 GACACCATCAACGAGAGAGGCGCGCGAGTGGACCGGTCGACCCCGTGACGCGCGGCC 660  
Db 598 GACACCATCAACGAGAGAGGCGCGCGAGTGGACCGGTCGACCCCGTGACGCGCGGCC 657  
QY 661 ATCGCGCGCGCGAGATGCGCGAGCGCCCGCGCGCAGCAGCATGCGCGCGCACCAAGCACC 720  
Db 658 GTGGCGCGCGCGCAGATGCGCGAGCGCCCGCGCGCAGCAGCATGCGCGCGCACCAAGCACC 717  
QY 721 CTGACAGAGAGATGCGCTGTGATGACCAAGACCCCGCATCCCGTGCGCGCATCTTAC 780  
Db 718 CTGACAGAGAGATGCGCTGTGATGACCAAGACCCCGCATCCCGTGCGCGCATCTTAC 777  
QY 781 AAGCGGTGATCATCTTGGGCTGAAACAAGATGTGCGGATGTACAGCCCGTGAGCATC 840  
Db 778 AAGCGGTGATCATCTTGGGCTGAAACAAGATGTGCGGATGTACAGCCCGTGAGCATC 837  
QY 841 CTGACATCAACGAGAGGCGCGCCCAAGAGCGCTTCCGCGCATGAGCGCTTCTCAAG 900  
Db 838 CTGACATCAACGAGAGGCGCGCCCAAGAGCGCTTCCGCGCATGAGCGCTTCTCAAG 897  
QY 901 ACCCTGCGCGCGCAGAGAGAGACCCAGAGAGTGAAGACTGATGACCGACCTGCTG 960  
Db 898 ACCCTGCGCGCGCAGAGAGAGACCCAGAGAGTGAAGACTGATGACCGACCTGCTG 957  
QY 961 GTGCAAGAGCGCGCAACCCCGCATGCAAGACCATCTTGGCGCTTCCGCGCGCGCGCAGC 1020  
Db 958 GTGCAAGAGCGCGCAACCCCGCATGCAAGACCATCTTGGCGCTTCCGCGCGCGCGCAGC 1017  
QY 1021 CTGAGAGAGATGATGACCGCTGCGCAGAGGCGCTGGCGCGCGCGCGCGCGCGCGCTG 1080  
Db 1018 CTGAGAGAGATGATGACCGCTGCGCAGAGGCGCTGGCGCGCGCGCGCGCGCGCGCTG 1077  
QY 1081 CTGGCGCGAGGCGATGAGCGCGCGCAACAC--AGCGTATGATGACAGAGAGCACTTC 1137  
Db 1078 CTGGCGCGAGGCGATGAGCGCGCGCAACAGCGTGAACATCATGATGACAGAGAGCACTTC 1137  
QY 1138 AAGGCGCGCGCGCGCATGCTGAAGTCTTCAACTGCGCGCAAGAGAGGCGCAATCGCCCGC 1197  
Db 1138 AAGGCGCGCGCGCGCAACGTAAGTCTTCAACTGCGCGCAAGAGAGGCGCAATCGCCCGC 1197  
QY 1198 AACTGCGCGCGCGCGCGCAAGAGGCGCTGTGAAGTGGCGCAAGAGGCGCACAGATG 1257  
Db 1198 AACTGCGCGCGCGCGCGCAAGAGGCGCTGTGAAGTGGCGCAAGAGGCGCACAGATG 1257  
QY 1258 AAGGACTGACCGAGCGCGCAGGCCAATTCTGGGCGAAGATTGGCCAGCCACAAGGGC 1317  
Db 1258 AAGGACTGACCGAGCGCGCAGGCCAATTCTGGGCGAAGATTGGCCAGCCACAAGGGC 1317  
QY 1318 CGCCCGCGCACTTCTGCAAGCG-----CCCGAGGCCACCGCC 1359  
Db 1318 CGCCCGCGCACTTCTGCAAGCGCGCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCC 1377  
QY 1360 CCCCCCGCGAGAGCTTCCGCTTGAGAGAGCAACCCCGCGCGAGAGAGAGAG 1419  
Db 1378 CCCCCCGCGAGAGCTTCCGCTTGAGAGAGCAACCCCGCGCGAGAGAGAGAG 1437  
QY 1420 G-----ACCGGAGACCTGACCACTGAGAGAGCTGTTGGCAAGACCC 1467  
Db 1438 GACCGGAGACCTTACCGGAGACCTGACCTGCGCAGCTGTGCGAGCGGCC 1497  
QY 1468 CTGAGCCAGTAA 1479  
Db 1498 CTGAGCCAGTAA 1509





```
QY 1 ATGGGCGCCGCGCAGCATCTGCGCGCGCAAGCTGGAACGCTGGAGCCATCCGC 60
Db 1880 ATGGCGCCGCGCGCAGCATCTGCGCGCGCAAGCTGGAACGCTGGAGCCATCAAG 1939
QY 61 CTGCGCCGCGCGCAAGAACTCTACATGATGAAGCAACCTGTGTGGGCGAGCCGCGAG 120
Db 1940 CTGAAGCCGCGCGCAAGAACTCTACATGATGAAGCAACCTGTGTGGGCGAGCCGCGAG 1999
QY 121 CTGGAAGATTGCGCCCTGGAACCCCGCGCTGTGAGAGCAAGCGAGGCTGCAAGCATC 180
Db 2000 CTGGAAGCTTGTGCGCCCTGGAACCCCGCGCTGTGAGAGCAAGCGAGGCTGCAAGCATC 2059
QY 181 ATCCGCGAGTGCACCCCGCGCTGAGACCGCGAGCGAGGCTGAAGAGCTGTTCAAC 240
Db 2060 ATGAAGCAGCTGCAGCCCGCGCTGAGACCGCGAGCGAGGCTGATCAGCTGCAACAAC 2119
QY 241 ACCGTGCGCAACCTGTACTGTGCGTGCACGAGAGATCGAGTCCGCGACACCAAGAGGCC 300
Db 2120 ACCGTGCGCAACCTGTACTGTGCGTGCACGAGAGATCGAGTCCGCGACACCAAGAGGCC 2179
QY 301 CTGGAACAAGTTCAGAGAGAGAGCAACAAGTCCAGAGAGATCCAGAGCGCGAGGCC 360
Db 2180 CTGGAACAAGTTCAGAGAGAGAGCAACAAGTCCAGAGAGATCCAGAGCGCGAGGCC 2239
QY 361 GCCGACAAGGCGCAAGGTGAGAGCCAGAACTACCCCATCTGTGCAAGAACTGCAAGGCCAGATG 420
Db 2240 GCCGACAAGGCGCAAGGTGAGAGCCAGAACTACCCCATCTGTGCAAGAACTGCAAGGCCAGATG 2299
QY 421 GTGCACCAAGCCATCAGCCCCCGCACCCCTGAAACGCTGTGTAAGTGTATCGAGAGAGAG 480
Db 2300 GTGCACCAAGCCATCAGCCCCCGCACCCCTGAAACGCTGTGTAAGTGTGTGAGAGAGAG 2359
QY 481 GCCTTCAGCCCCGAGGTGATTCCTCATGTTCAACCGCCCTGAGCGAGGCGCCACCCCGCAG 540
Db 2360 GCCTTCAGCCCCGAGGTGATTCCTCATGTTCAACCGCCCTGAGCGAGGCGCCACCCCGCAG 2419
QY 541 GACCTGAACAAGATGTTGAACAACGCTGCGCGCGCCACCAAGGCCCATGCAAGTCTGAAG 600
Db 2420 GACCTGAACAAGATGTTGAACAACGCTGCGCGCGCCACCAAGGCCCATGCAAGTCTGAAG 2479
QY 601 GACACCATCAACGAGAGAGGCGCGCGAGTGGAGACCGCGTGCACCCCGTGCACCGCGGCC 660
Db 2480 GACACCATCAACGAGAGAGGCGCGCGAGTGGAGACCGCGTGCACCCCGTGCACCGCGGCC 2539
QY 661 ATGCGCCCCCGCGCGAGATGCGCGAGAGCCCGCGCGCGAGCAATCGCCGCGCACCAACCAAGCAC 720
Db 2540 ATGCGCCCCCGCGCGAGATGCGCGAGAGCCCGCGCGCGAGCAATCGCCGCGCACCAACCAAGCAC 2599
QY 721 CTGCAGAGCAAGATGCGCTGTGATGACCAAGCAACCCCGCATTCCTGTGGCGCATCTAC 780
Db 2600 CTGCAGAGCAAGATGCGCTGTGATGACCAAGCAACCCCGCATTCCTGTGGCGCATCTAC 2659
QY 781 AAGCGGTGATCATCTGTGGCGCTGAACAAGATGCTGCGGATGTACAGCGCGGTGAGCATC 840
Db 2660 AAGCGGTGATCATCTGTGGCGCTGAACAAGATGCTGCGGATGTACAGCGCGGTGAGCATC 2719
QY 841 CTGCACATCAAGCAGGCGCCCAAGAGACCTTCCGCGACTACGTGAGCCGCTTCTTCAAG 900
Db 2720 CTGCACATCAAGCAGGCGCCCAAGAGACCTTCCGCGACTACGTGAGCCGCTTCTTCAAG 2779
QY 901 ACCCTGCGCGCGCGAGCAGAGCAACCAAGAGGTGAAGAACTGAGATGACCGACACCCCTGTG 960
Db 2780 ACCCTGCGCGCGCGAGCAGAGCAACCAAGAGGTGAAGAACTGAGATGACCGACACCCCTGTG 2839
QY 961 GTGCAGAAAGCCCAACCCCGCACTGCAAGACCATCTGCGCGCTCTCGGCGCGCGCGCGAGC 1020
Db 2840 GTGCAGAAAGCCCAACCCCGCACTGCAAGACCATCTGCGCGCTCTCGGCGCGCGCGCGAGC 2899
QY 1021 CTGGAAGAGATGATGACCGCTGCGCAGAGGCGTGGCGCGCGCGCGCGCGCGCGCGCTG 1080
Db 2900 CTGGAAGAGATGATGACCGCTGCGCAGAGGCGTGGCGCGCGCGCGCGCGCGCGCGCTG 2959
```

```
QY 1081 CTGGCCGAGGCGATGAGCCAGG---CCAACAACGCGTGTATGATGAGAAGACAATTTC 1137
Db 2960 CTGGCCGAGGCGATGAGCAAGGTGAACAACAACCATATGATGACAGCGCAACTGC 3019
QY 1138 AAGGGCCCCCGCGCATCTGTCAAGTGTCTTCAACTGCGGCAAGAGGGGCCCATCGCCCGC 1197
Db 3020 AAGGGCCCCCAAGCGCAACCATCAAGTGTCTTCAACTGCGGCAAGAGGGGCCCATCGCCCGC 3079
QY 1198 AACTCGCGCGCGCGCGCAAGAGGGCTGTGTAAGTGTGCGGCAAGAGGGGCCCAAGATG 1257
Db 3080 AACTCGCGCGCGCGCGCAAGAGGGCTGTGTAAGTGTGCGGCAAGAGGGGCCCAAGATG 3139
QY 1258 AAGGACTGCACCGAGCGCGCAAGCCACTTCTGCGGCAAGATGTGCGCCCAAGCAAGGGC 1317
Db 3140 AAGGACTGCACCGAGCGCGCAAGCCACTTCTGCGGCAAGATGTGCGCCCAAGCAAGGGG 3194
QY 1318 CGCCCGCGCAACTTCTGTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 1377
Db 3195 AGGCCAGGGAATTTCTTCAAGAGCAGACCGAGCGCAACAGCGCCCAAGAGAGAGCTTC 3254
QY 1378 CGCTTCGAGGAG-----GACCAACCCCGCGCGCAAGCAGAG-----AGCAAGAGCGCG 1425
Db 3255 AGGTTGGGGAAGAGACAACAACCTCTCTGAGAGCAGAGAGCGGATAGACAGAGAACTG 3314
QY 1426 GAGACCCCTGACCAAGCCTGGAAGAGCGCTGTGCGCAACGACCCCTGAGCCAGTAA 1479
Db 3315 TATCCTTAGCTTCCTCAGATCACTTGTGCGAGCGAGCCCTGTCACAATTA 3368

RESULT 14
AX149648
LOCUS AX149648 4288 bp DNA linear PAT 08-JUN-2001
DEFINITION Sequence 2 from Patent WO0136614.
ACCESSION AX149648
VERSION AX149648.1 GI:14348047
KEYWORDS
SOURCE
ORGANISM Human immunodeficiency virus
Human immunodeficiency virus
Viruses; Retrovird viruses; Retroviridae; Lentivirus; Primate
lentivirus group.
REFERENCE
1 Shao, Y., Wagner, R., Wolf, H. and Graf, M.
The genome of the hiv-1 inter-subtype (c/b') and use thereof
Patent: WO 0136614-A 2 25-MAY-2001;
Geneart GMBH Gesellschaft fuer angewandte Biotechnologie (DE) ;
Shao, Yiming (CN)
FEATURES
Location/Qualifiers
source 1..4288
/organism="Human immunodeficiency virus"
/mol_type="unassigned DNA"
/db_xref="taxon:12721"
ORIGIN
Query Match 81.5%; Score 1205.4; DB 6; Length 4288;
Best Local Similarity 89.5%; Pred. No. 1.1e-121;
Matches 1335; Conservative 0; Mismatches 141; Indels 15; Gaps 3;

QY 1 ATGGGCGCGCGCGCGCAAGCATCTGCGCGCGCAAGCTGGAACGCTGGAGCCATCCGC 60
Db 13 ATGGGCGCGCGCGCGCAAGCATCTGAGGCGCGCAAGCTGGAACAAGTGGAGAAATCAGG 72
QY 61 CTGCGCGCGCGCGCGCAAGAGTGTCTACATGATGAAGCAACCTGTGTGGGCGAGCCGCGAG 120
Db 73 CTGAGGCGCGCGCGCGCAAGAGCACTACATGCTGAAGCACTGTGTGTGGGCGAGGAG 132
QY 121 CTGGAAGATTGCGCCCTGGAACCCCGCGCTGTGAGAGCAAGCGAGGCTGCAAGCAGATC 180
Db 133 CTGAAGAGGTTGCGCTTGAACCCCGCGCTGTGAGAGCAAGCGAGGCTGCAAGCAGATC 192
QY 181 ATCCGCGAGTGCACCCCGCGCTGCAAGCGCGCGCGCGCGCGCGCGCGCTGTTCAAC 240
Db 193 ATGAAGCAGCTGCAGAGCGCGCTGCAAGCGCGCGCGCGCGCGCGCGCGCTGTTCAAC 252
```



QY	241	ACCGTGGCCACCTGTACTGTCGTGCA	CGAGAAATCGAGTCCGCGACACCAAGAGGCC	300	
Db	253	ACCGTGGCCACCCCTACTGTGTGCA	CACCGAATCGACGTGAGGACACCAAGAGGCC	312	
QY	301	CTGGAACAAGATCGAGGAGGACAGAA	CAAGTCCAGCAGAAATCCAGAGGCCGAGGCC	360	
Db	313	CTGGAACAAGATCGAGGAGGAGCAGAA	CAAGATCCAGCAGAAAGCACCAGAGGCCAAGAG	372	
QY	361	GCCGACAAGGGCAAGGTGAGCCAGAA	CTAACCCATGTGCAAACTTGAGGGCCAGATG	420	
Db	373	GCCGAC---GGCAAGGTGAGCCAGAA	CTAACCCATGTGCAAACTTGAGGGCCAGATG	429	
QY	421	GTGACACAGGCCCATCAGCCCCGCG	ACCTGAACGCTGGGTGAAGTATCGAGAGAG	480	
Db	430	GTGACACAGGCCCATCAGCCCCAGAA	CCCTGAATGCATGGGTGAAGTGTGAGAGAG	489	
QY	481	GCCTTCAGCCCCGAGGTGATCCCAT	GTTCACCGCCTGAGCGAGGGCCACCCCCAG	540	
Db	490	GCCTTCAGCCCCGAGGTGATCCCAT	GTTCAGCGCCTGAGCGAGGGCCACCCCCAG	549	
QY	541	GACCTGAACA	CGATGTTGAACA	CGTGGCGGCCACAGGCCGCATGCA	600
Db	550	GACCTGAACA	CGATGTTGAACA	CGTGGCGGCCACAGGCCGCATGCA	609
QY	601	GACACCATCAACGAGGAGCGCGAG	TGGGACCGCGTGCA	CCCGTGCA	660
Db	610	GACACCATCAACGAGGAGCGCGAG	TGGGACCGCGTGCA	CCCGTGCA	669
QY	661	ATGCCCCCGCCAGATGCGCGAG	CCCGCGCGACATCGCGGCCACCA	CAGCAAC	720
Db	670	ATGCCCCCGCCAGATGCGCGAG	CCCGCGCGACATCGCGGCCACCA	CAGCAAC	729
QY	721	CTGACAGACAGATCGCTGTGATGA	CCAGCAACCCCCCATCCCGTGGCGACATTA	C	780
Db	730	CTGACAGACAGATCGCTGTGATGA	CCAGCAACCCCCCATCCCGTGGCGACATTA	C	789
QY	781	AAGCGGTGATCATCTGGGCTTGA	CAAGATGTGCGANTGTACGCCCGTGACATC	840	
Db	790	AAGAGGTGATCATCTGGGCTTGA	CAAGATGTGAGATGTACGCCCCAGCAGATC	849	
QY	841	CTGACATCAAGCAGGGCCCCCAAG	AGGCCCTTCGCGACTACGTGACCGCTTCAAG	900	
Db	850	CTGACATCAAGCAGGGCCCCCAAG	AGGCCCTTCGCGACTACGTGACCGCTTCAAG	909	
QY	901	ACCTGCGCGCCGAGCAGAGCA	CCAGAGGTGAAGAACTGATGACCGACACCTGCTG	960	
Db	910	ACCTGAGGGCGGAGCAGGCCAC	CCAGGGCGTGAAGAACTGATGACCGACACCTGCTG	969	
QY	961	GTGACAGACGCCAACCCCGACTG	CAAGACCATCTGCGCGCTTCGCGCCCCGCGCCAGC	1020	
Db	970	GTGACAGACGCCAACCCCGACTG	CAAGACCATCTGAGGGCCCTTGGGCCCCGCGCCAGC	1029	
QY	1021	CTGAGAGATGATGACCGCTG	CCAGGGCGGTGGCGGCCCCAGACCA	MAAGGCCCGCGTG	1080
Db	1030	ATCGAGAGATGATGACCGCTG	CCAGGGCGGTGGCGGCCCCAGACCA	MAAGGGCCAGGTG	1089
QY	1081	CTGGCCGAGGCGATGAGCCAGG	CAACACAGCCGTGATGATGCAAGAGCAACTTCAAG	1140	
Db	1090	CTGGCCGAGGCGATGAGCCAGAG	CAACACAGCCGTGATGCAAGAGAGCAACTTCAAG	1149	
QY	1141	GGCCCCCGGCGCATGTCAGATG	CTTCAACTGCGGCAAGAGGGCCACATCGCCGCAAC	1200	
Db	1150	GGCAGCAAGAGATGTCGTAAGT	CTTCAACTGCGGCAAGAGGGCCACATCGCCGCAAC	1209	
QY	1201	TGCGCGCGCCCGCCAGAGAGGG	CTGTGAAGTGGCCAGAGAGGGCCACAGATGAAG	1260	
Db	1210	TGCAAGGCGCCCGCCAGAGAGGG	CTGTGAAGTGGCCAGAGAGGGCCACAGATGAAG	1269	
QY	1261	GACTGCACCGAGCGCCAGG	CCAACTTCTGTGGCAAGATCTGGCCAGCCCA	MAAGGGCGGC	1320
Db	1270	GACTGCACCGAGAGGCAAG	CCAACTTCTGTGGCAAGATCTGGCCAGCCCA	MAAGGGCGGC	1329
QY	1321	CCCGGCAACTTCTGCAAGAG	CGCGCCCGAGCCACAGCGCCCCCCCCCGCGAGAGCTTCCGC	1380	

```

Db      1330 CCGGCAACTTCTGTGAGAACAGAGCCCGAGCCACCGCCCCCGAGAGAGCTTCAGG 1389
        |||
QY      1381 TTGGAGGAG-----ACCACCCCGCGCGCAAGACGAG-----AGCAAGACCGCGAG 1428
        |||
Db      1390 TTGGAGGAGGAGAGACCAACCCCGAGCGAAGACGAGCCCATGACAAAGAGCTGTAC 1449
        |||
QY      1429 ACCGTGACCGAGCTGGAAGAGCTGTTCGGCAACGAGCCCTGTGAGCCAGTAA 1479
        |||
Db      1450 CCCGTGACCGAGCTGGAAGAGCTGTTCGGCAACGAGCCCGAGCGACGAGAA 1500
        |||

RESULT 15
LOCUS   AF201927                      1509 bp    DNA        linear      SYN 16-MAR-2000
DEFINITION Synthetic construct gag protein gene, complete cds.
ACCESSION AF201927
VERSION   AF201927.1  GI:7248702
KEYWORDS
SOURCE    '
ORGANISM  synthetic construct
          synthetic construct
          artificial sequences.
REFERENCE 1 (bases 1 to 1509)
AUTHORS   zur Megede,J., Chen,M.C., Doe,B., Schaefer,M., Greer,C.E.,
          Selby,M., Otten,G.R. and Barnett,S.W.
TITLE     Increased expression and immunogenicity of sequence-modified human
          immunodeficiency virus type 1 gag gene
JOURNAL   J. Virol. 74 (6), 2628-2635 (2000)
MEDLINE   20148954
PUBMED    10684277
REFERENCE 2 (bases 1 to 1509)
AUTHORS   zur Megede,J. and Barnett,S.W.
TITLE     Direct Submission
JOURNAL   Submitted (04-NOV-1999) Vaccines, Chiron Corporation, 4560 Horton,
          Emeryville, CA 94608, USA

FEATURES
Source    location/Qualifiers
          1..1509
             /organism="synthetic construct"
             /mol_type="genomic DNA"
             /db_xref="taxon:32630"
             1..1509
             /note="HIV-1SF2 p55gag"
             /codon_start=1
             /transl_table=1
             /product="gag protein"
             /protein_id="AAF43628.1"
             /db_xref="GI:7248703"
             /translation="MGARASVLSGGLDKWEKIRLRDGGKKYKLKHIVASRELERF
             AVNPGILLETSEGRQILGLQPSLOTGSBELRSLVNTVATLYCVHORIDVDTKALB
             KIEBKONKSKKKAQOAAAAGTGNSSOVSNPIVONLQGMVHOALSPRTLNAVVKV
             VEEKAPSPVAPMFSAISSEGTADITSTLQEOIGMWTNPPIPVGBIYKRWITLGLNKIV
             PVHAGPIAPQMRBPRGSDIAGTITLQEOIGMWTNPPIPVGBIYKRWITLGLNKIV
             RMYSPSTILDRQGPKEPRFDYDRFYKTLRAEQASQDVKNWMTETLLVQNPNDCKT
             ILKALGPATLEEMTTACQGVGGPHKARVLAEMSQTNPATIMORGNFRNQKTV
             KCFNGKEGHTARNCRAPRKKGCRGREGHQMDCTEROANFLGIWPSYKGRPGNF
             LQSRPEPTAPBESFRFGKEKTPSQKQEPIDKELFPLTSLRSLFGNDPSSQ"

CDS
          1..1509
             /note="HIV-1SF2 p55gag"
             /codon_start=1
             /transl_table=1
             /product="gag protein"
             /protein_id="AAF43628.1"
             /db_xref="GI:7248703"

ORIGIN
Query Match      81.0%; Score 1197.8; DB 12; Length 1509;
Best Local Similarity 89.3%; Pred. No. 9,9e-121;
Matches 1347; Conservative 0; Mismatches 132; Indels 30; Gaps 4;
QY      1 ATGGCGCGCGCGCGCGCAGCATCTGCGCGCGCGCGCAAGCTGAGCGCTCGGAGCGCATCCG 60
        |||
Db      1 ATGGCGCGCGCGCGCGCAGCGCTGTGAGCGCGCGCGAGCTGGAACAAGTGGAGAAGATCCGC 60
        |||
QY      61 CTGCGCGCGCGCGCGCGCAAGAAGTGTCTACATGATGAAGCACCTGTGTTCGGCCAGCGCGAG 120
        |||
Db      61 CTGCGCGCGCGCGCGCGCAAGAAGTACAAAGCTGAAGCACATCGTGTTCGGCCAGCGCGAG 120
        |||
QY      121 CTGGAAGAAGTTGCGCCCTGAACCCCGGCTGTCTGGAAGACCAAGCGAGGGCTGCAAGCAGATC 180
        |||
Db      121 CTGGAAGCGCTTGCOCGTGAACCCCGGCTGTCTGGAAGACCAAGCGAGGGCTGCGCCAGATC 180
        |||

```



QY 181 ATCCGCGAGTGCACCCCGCCTGCGAGACCGGAGAGCTGAAGAGCTGTTCAAC 240  
DB 181 CTGGGCGCAGCTGCAGCCCGCTGCGAGACCGGAGAGAGCTGCGCAGCTGTACAAC 240  
QY 241 ACCGTGGCCACCTGTGTACTGCTGTCAGAGAAATCGAGGTCCGCGACCAAGAGGCC 300  
DB 241 ACCGTGGCCACCTGTGTACTGCTGTCAGAGCAATCGTCAAGACACCAAGAGGCC 300  
QY 301 CTGGAAGAATCGAGAGGAGCAACAAGTCCAGAGAAATCGAGAGCCAGCGCGGCC 360  
DB 301 CTGGAAGAATCGAGAGGAGCAACAAGTCCAGAGAAAGGCCAGCGCGGCC 360  
QY 361 GCGG-----ACAAGGCGAAGTGAAGCCGCACTACCCCAATGTCAGAACCTG 408  
DB 361 GCGGCGCGGCAACCGGCAACAGCAGGAGTGAAGCCGCACTACCCCAATGTCAGAACCTG 420  
QY 409 CAGGCGCCAGATGCTGTCACAGGCGCATCAGCCCGGCAACCTGAACGCTGGTGAAGTG 468  
DB 421 CAGGCGCCAGATGCTGTCACAGGCGCATCAGCCCGGCAACCTGAACGCTGGTGAAGTG 480  
QY 469 ATCGAGGAGAAAGGCTTTCAGGCGCGAGGTGATCCCCATGTTTCAGCGGCTGAGCGAGGGC 528  
DB 481 GTGAGGAGAAAGGCTTTCAGGCGCGAGGTGATCCCCATGTTTCAGCGGCTGAGCGAGGGC 540  
QY 529 GCCACCGCCCGCAGGACCTGAACACCATGTTGAACACCGTGGCGGCCACAGCGCGCATG 588  
DB 541 GCCACCGCCCGCAGGACCTGAACACCATGTTGAACACCGTGGCGGCCACAGCGCGCATG 600  
QY 589 CAGATGCTGAAGGACACCATCAACGAGAGGCGCGGAGTGGGACCGCGTGCACCCCGTG 648  
DB 601 CAGATGCTGAAGGAGACCATCAACGAGAGGCGCGGAGTGGGACCGCGTGCACCCCGTG 660  
QY 649 CAGCGCGCGCCCATGCGCGCGCGCGCGAGTGCAGCGCGCGCGCGCGCGCATGCGCGGC 708  
DB 661 CAGCGCGCGCCCATGCGCGCGCGCGCGAGTGCAGCGCGCGCGCGCGCGCATGCGCGGC 720  
QY 709 ACCACCGAGCACCTGTCAGAGGAGCATGCGCTGATGAACAGCAACCCCGCATCCCGTG 768  
DB 721 ACCACCGAGCACCTGTCAGAGGAGCATGCGCTGATGAACAGCAACCCCGCATCCCGTG 780  
QY 769 GCGGACATCTACAGCGGTGATCATCTGGGCTGAACAAGATGTCGGATGTACAGC 828  
DB 781 GCGGAGATCTACAGCGGTGATCATCTGGGCTGAACAAGATGTCGGATGTACAGC 840  
QY 829 CCGCTGAGCATCTGGAATCAAGGAGCGCGCGCGCGCGCGCGCGCGCATGCGTGAC 888  
DB 841 CCGACCGAGCATCTGGAATCAAGGAGCGCGCGCGCGCGCGCGCGCGCATGCGTGAC 900  
QY 889 CGCTTCTTCAAGACCTTGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 948  
DB 901 CGCTTCTTCAAGACCTTGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 960  
QY 949 GACACCTGCTGTCAGAAACCGCAACCCCGAGTGCAGAACCATCTGCGCGCTTCGCGC 1008  
DB 961 GAGACCTGCTGTCAGAAACCGCAACCCCGAGTGCAGAACCATCTGCGCGCTTCGCGC 1020  
QY 1009 CCGGCGCGCAGCTGAGAGAGATGACCGGCTGCGAGGCGGTGGCGCGCGCGCGCGCGC 1068  
DB 1021 CCGGCGCGCAGCTGAGAGAGATGACCGGCTGCGAGGCGGTGGCGCGCGCGCGCGCGC 1080  
QY 1069 AAGGCGCGGCTGCTGCGAGGCGGATGAGCCAGGTGACGAACCGGCGAGCATATGATG 1122  
DB 1081 AAGGCGCGGCTGCTGCGAGGCGGATGAGCCAGGTGACGAACCGGCGAGCATATGATG 1140  
QY 1123 CAGAGAAGCACTTCAAGGCGCGCGCGCGCGCATGCTCAAGTCTTCAACTGCGGCAAGAG 1182  
DB 1141 CAGGCGGCAACTTCCGCAACCGGAGAGACCGTCAAGTCTTCAACTGCGGCAAGAG 1200  
QY 1183 GGCACATGCGCGCAACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1242  
DB 1201 GGCACATGCGCGCAACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1260

QY 1243 GAGGCGCACCGAGATGAAGGACTGCAACCGAGCGCGGCGGCACTTCTTGGGCAAGATCTG 1302  
DB 1261 GAGGCGCACCGAGATGAAGGACTGCAACCGAGCGCGGCGGCACTTCTTGGGCAAGATCTG 1320  
QY 1303 CCGAGCCCAAGGCGCGCGCGCGGCACTTCTTGAAGAGCGCGCGCGCGCGCGCGCGCG 1362  
DB 1321 CCGAGCTTCAAGGCGCGCGCGGCACTTCTTGAAGAGCGCGCGCGCGCGCGCGCGCG 1380  
QY 1363 CCGGCGGAGAGCTTCCGCTTC-----GAGGAGACCAACCGCGCGGCAAGAGAGAG--- 1413  
DB 1414 ---AGCAAGGAGCGCGGAGACCTGACCAAGCTGTAAGAGCTGTTGGCAACGACCCCTG 1470  
QY 1441 ATGCAAGAGAGCTGTACCCCTGACCAAGCTGCGGAGCTGTTGGCAACGACCCGAGC 1500  
DB 1471 AGCGAGTAA 1479  
DB 1501 AGCGAGTAA 1509

Search completed: May 28, 2004, 13:23:24  
Job time : 6085.87 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 28, 2004, 09:24:06 ; Search time 6.49156 Seconds  
(without alignments)  
5129.286 Million cell updates/sec

Title: US-09-475-704A-2

Perfect score: 60  
Sequence: 1 gacatccgcagggcccccacaa.....tgacacgcttctcaagacc 60

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 682709 segs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA: \*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq: \*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq: \*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq: \*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq: \*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq: \*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	58.4	97.3	60	US-09-475-515-20	Sequence 20, Appli
2	58.4	97.3	1268	US-09-475-515-9	Sequence 9, Appli
3	58.4	97.3	1515	US-09-475-515-4	Sequence 4, Appli
4	58.4	97.3	1853	US-09-475-515-5	Sequence 5, Appli
5	58.4	97.3	1865	US-09-475-515-78	Sequence 78, Appli
6	58.4	97.3	1865	US-09-475-515-79	Sequence 79, Appli
7	58.4	97.3	2031	US-09-475-515-7	Sequence 7, Appli
8	58.4	97.3	4319	US-09-475-515-6	Sequence 6, Appli
9	58.4	97.3	4472	US-09-475-515-75	Sequence 75, Appli
10	58.4	97.3	4608	US-09-475-515-76	Sequence 76, Appli
11	58.4	97.3	4689	US-09-475-515-74	Sequence 74, Appli
12	58.4	97.3	4766	US-09-475-515-73	Sequence 73, Appli
13	47.8	79.7	4307	US-09-552-950-2	Sequence 2, Appli
14	47.8	79.7	9772	US-09-552-950-5	Sequence 5, Appli
15	35	58.3	1476	US-09-184-418C-100	Sequence 100, App
16	35	58.3	1485	US-09-184-418C-73	Sequence 73, Appli
17	35	58.3	8959	US-09-184-418C-11	Sequence 11, Appli
18	35	58.3	9010	US-09-184-418C-8	Sequence 8, Appli
19	33.4	55.7	718	5204259-2	Patent No. 5204259
20	33.4	55.7	793	5204259-4	Patent No. 5204259
21	33.4	55.7	796	5204259-6	Patent No. 5204259
22	33.4	55.7	845	US-08-589-446-3	Sequence 3, Appli
23	33.4	55.7	845	US-08-444-882-3	Sequence 3, Appli
24	33.4	55.7	845	US-08-389-459A-3	Sequence 3, Appli
25	33.4	55.7	845	US-08-987-867A-3	Sequence 3, Appli
26	33.4	55.7	871	5204259-10	Patent No. 5204259
27	33.4	55.7	1021	5204259-8	Patent No. 5204259

28	33.4	55.7	4307	4	US-09-552-950-1	Sequence 1, Appli
29	33.4	55.7	4338	4	US-09-872-733A-1	Sequence 1, Appli
30	33.4	55.7	5362	3	US-08-463-210-5	Sequence 5, Appli
31	33.4	55.7	5362	4	US-08-463-028-5	Sequence 5, Appli
32	33.4	55.7	7228	2	US-08-850-049-128	Sequence 128, App
33	33.4	55.7	7228	2	US-08-850-049-129	Sequence 128, App
34	33.4	55.7	7228	2	US-08-050-478-128	Sequence 128, App
35	33.4	55.7	7228	2	US-08-050-478-129	Sequence 129, App
36	33.4	55.7	7228	3	US-09-414-117-128	Sequence 128, App
37	33.4	55.7	7228	3	US-09-414-117-129	Sequence 129, App
38	33.4	55.7	7228	4	US-09-678-437-128	Sequence 128, App
39	33.4	55.7	7228	4	US-09-678-437-129	Sequence 129, App
40	33.4	55.7	7399	2	US-08-418-848A-9	Sequence 9, Appli
41	33.4	55.7	8366	4	US-09-872-733A-6	Sequence 6, Appli
42	33.4	55.7	9709	2	US-08-188-583-5	Sequence 5, Appli
43	33.4	55.7	9709	3	US-08-388-353-1	Sequence 1, Appli
44	33.4	55.7	9709	3	US-08-488-551B-1	Sequence 1, Appli
45	33.4	55.7	9709	4	US-09-309-572-15	Sequence 15, Appli

## ALIGNMENTS

RESULT 1  
US-09-475-515-20  
Sequence 20, Application US/09475515A  
Patent No. 6602705  
GENERAL INFORMATION:  
APPLICANT: BARNETT, Susan  
APPLICANT: ZUR MEGEDR, Jan  
APPLICANT: SRIVASTAVA, Indresh  
APPLICANT: LIAN, Ying  
APPLICANT: HARTOG, Karin  
APPLICANT: GREER, Catherine  
APPLICANT: SELBY, Mark  
APPLICANT: WALKER, Christopher  
TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION  
FILE REFERENCE: 1621.002  
CURRENT APPLICATION NUMBER: US/09/475,515A  
CURRENT FILING DATE: 1999-12-30  
NUMBER OF SEQ ID NOS: 90  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 20  
LENGTH: 60  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic p55  
US-09-475-515-20

Query Match 97.3% Score 58.4; DB 4; Length 60;  
Best Local Similarity 98.3% Pred. No. 2.7e-10;  
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GACATCCGCGAGGGCCCAAGAGACCTTCGCGACTACGTGACCGCTTCTCAAGACC 60  
DB 1 GACATCCGCGAGGGCCCAAGAGACCTTCGCGACTACGTGACCGCTTCTCAAGACC 60

RESULT 2  
US-09-475-515-9  
Sequence 9, Application US/09475515A  
Patent No. 6602705  
GENERAL INFORMATION:  
APPLICANT: BARNETT, Susan  
APPLICANT: ZUR MEGEDR, Jan  
APPLICANT: SRIVASTAVA, Indresh  
APPLICANT: LIAN, Ying  
APPLICANT: HARTOG, Karin  
APPLICANT: LIU, Hong

APPLICANT: GREER, Catherine  
APPLICANT: SELBY, Mark  
TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION  
FILE REFERENCE: 1621.002  
CURRENT APPLICATION NUMBER: US/09/475,515A  
NUMBER OF SEQ ID NOS: 90  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 9  
LENGTH: 1268  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic Gag  
US-09-475-515-9

Query Match 97.3%; Score 58.4; DB 4; Length 1268;  
Best Local Similarity 98.3%; Pred. No. 3.5e-10;  
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GACATCCGCCAGGCCCCCAAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 60  
DB 862 GACATCCGCCAGGCCCCCAAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 921

RESULT 3  
US-09-475-515-4  
Sequence 4, Application US/09475515A  
Patent No. 6602705

GENERAL INFORMATION:  
APPLICANT: BARNETT, Susan  
APPLICANT: ZUR MEGEDE, Jan  
APPLICANT: SRIVASTAVA, Indresh  
APPLICANT: LIAN, Ying  
APPLICANT: HARTOG, Karin  
APPLICANT: LIU, Hong  
APPLICANT: GREER, Catherine  
APPLICANT: SELBY, Mark  
APPLICANT: WALKER, Christopher  
TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION  
FILE REFERENCE: 1621.002  
CURRENT APPLICATION NUMBER: US/09/475,515A  
NUMBER OF SEQ ID NOS: 90  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 4  
LENGTH: 1515  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-09-475-515-4

Query Match 97.3%; Score 58.4; DB 4; Length 1515;  
Best Local Similarity 98.3%; Pred. No. 3.6e-10;  
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GACATCCGCCAGGCCCCCAAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 60  
DB 862 GACATCCGCCAGGCCCCCAAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 921

RESULT 4  
US-09-475-515-5  
Sequence 5, Application US/09475515A  
Patent No. 6602705  
GENERAL INFORMATION:  
APPLICANT: BARNETT, Susan

APPLICANT: ZUR MEGEDE, Jan  
APPLICANT: SRIVASTAVA, Indresh  
APPLICANT: LIAN, Ying  
APPLICANT: HARTOG, Karin  
APPLICANT: LIU, Hong  
APPLICANT: GREER, Catherine  
APPLICANT: SELBY, Mark  
APPLICANT: WALKER, Christopher  
TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION  
FILE REFERENCE: 1621.002  
CURRENT APPLICATION NUMBER: US/09/475,515A  
NUMBER OF SEQ ID NOS: 90  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 5  
LENGTH: 1853  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-09-475-515-5

Query Match 97.3%; Score 58.4; DB 4; Length 1853;  
Best Local Similarity 98.3%; Pred. No. 3.7e-10;  
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GACATCCGCCAGGCCCCCAAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 60  
DB 862 GACATCCGCCAGGCCCCCAAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 921

RESULT 5  
US-09-475-515-78  
Sequence 78, Application US/09475515A  
Patent No. 6602705

GENERAL INFORMATION:  
APPLICANT: BARNETT, Susan  
APPLICANT: ZUR MEGEDE, Jan  
APPLICANT: SRIVASTAVA, Indresh  
APPLICANT: LIAN, Ying  
APPLICANT: HARTOG, Karin  
APPLICANT: LIU, Hong  
APPLICANT: GREER, Catherine  
APPLICANT: SELBY, Mark  
APPLICANT: WALKER, Christopher  
TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION  
FILE REFERENCE: 1621.002  
CURRENT APPLICATION NUMBER: US/09/475,515A  
NUMBER OF SEQ ID NOS: 90  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 78  
LENGTH: 1865  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: GPI  
US-09-475-515-78

Query Match 97.3%; Score 58.4; DB 4; Length 1865;  
Best Local Similarity 98.3%; Pred. No. 3.7e-10;  
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GACATCCGCCAGGCCCCCAAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 60  
DB 868 GACATCCGCCAGGCCCCCAAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 927

RESULT 6  
US-09-475-515-79

```
; Sequence 79, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 79
; LENGTH: 1865
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GP2
US-09-475-515-79
```

```
Query Match          97.3%; Score 58.4; DB 4; Length 1865;
Best Local Similarity 98.3%; Pred. No. 3.7e-10;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY      1 GACATCCGCGCAGGGCCCCCAAGAGACCCCTTCCGCGACTAGTGACCGCTTCTCAAGACC 60
DB      868 GACATCCGCGCAGGGCCCCCAAGAGACCCCTTCCGCGACTAGTGACCGCTTCTCAAGACC 927
```

## RESULT 7

```
US-09-475-515-7
; Sequence 7, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 2031
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: HIV-Gag/HCV-core fusion polypeptide
US-09-475-515-7
```

```
Query Match          97.3%; Score 58.4; DB 4; Length 2031;
Best Local Similarity 98.3%; Pred. No. 3.7e-10;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY      1 GACATCCGCGCAGGGCCCCCAAGAGACCCCTTCCGCGACTAGTGACCGCTTCTCAAGACC 60
DB      862 GACATCCGCGCAGGGCCCCCAAGAGACCCCTTCCGCGACTAGTGACCGCTTCTCAAGACC 921
```

```
RESULT 8
US-09-475-515-6
; Sequence 6, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 4319
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: HIV-Gag-polymerase
US-09-475-515-6
```

```
Query Match          97.3%; Score 58.4; DB 4; Length 4319;
Best Local Similarity 98.3%; Pred. No. 4e-10;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY      1 GACATCCGCGCAGGGCCCCCAAGAGACCCCTTCCGCGACTAGTGACCGCTTCTCAAGACC 60
DB      862 GACATCCGCGCAGGGCCCCCAAGAGACCCCTTCCGCGACTAGTGACCGCTTCTCAAGACC 921
```

## RESULT 9

```
US-09-475-515-75
; Sequence 75, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 75
; LENGTH: 4472
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: gp160.modUS4.delV1/V2.gag.modSF2
US-09-475-515-75
```

```
Query Match          97.3%; Score 58.4; DB 4; Length 4472;
Best Local Similarity 98.3%; Pred. No. 4e-10;
```

Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GACATCCGCCAGGGCCCCCAAGAGCCCTTCGCGACTAGTGACCGCTTCTTCAAGACC 60  
Db 3794 GACATCCGCCAGGGCCCCCAAGAGCCCTTCGCGACTAGTGACCGCTTCTTCAAGACC 3853

## RESULT 10

US-09-475-515-76

; Sequence 76, Application US/09475515A

; Patent No. 6602705

; GENERAL INFORMATION:

; APPLICANT: BARNETT, Susan

; APPLICANT: ZUR MEGEDE, Jan

; APPLICANT: SRIVASTAVA, Indresh

; APPLICANT: LIAN, Ying

; APPLICANT: HARTOG, Karin

; APPLICANT: LIU, Hong

; APPLICANT: GREER, Catherine

; APPLICANT: SELBY, Mark

; APPLICANT: WALKER, Christopher

; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION

; FILE REFERENCE: 1621.002

; CURRENT FILING DATE: 1999-12-30

; NUMBER OF SEQ ID NOS: 90

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 76

; LENGTH: 4608

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence:

; OTHER INFORMATION: gp160.modsF162.delV2.gag.modsF2

; US-09-475-515-76

; US-09-475-515-76

Query Match 97.3%; Score 58.4; DB 4; Length 4608;

Best Local Similarity 98.3%; Pred. No. 4e-10;

Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GACATCCGCCAGGGCCCCCAAGAGCCCTTCGCGACTAGTGACCGCTTCTTCAAGACC 60  
Db 3930 GACATCCGCCAGGGCCCCCAAGAGCCCTTCGCGACTAGTGACCGCTTCTTCAAGACC 3989

## RESULT 11

US-09-475-515-74

; Sequence 74, Application US/09475515A

; Patent No. 6602705

; GENERAL INFORMATION:

; APPLICANT: BARNETT, Susan

; APPLICANT: ZUR MEGEDE, Jan

; APPLICANT: SRIVASTAVA, Indresh

; APPLICANT: LIAN, Ying

; APPLICANT: HARTOG, Karin

; APPLICANT: LIU, Hong

; APPLICANT: GREER, Catherine

; APPLICANT: SELBY, Mark

; APPLICANT: WALKER, Christopher

; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION

; FILE REFERENCE: 1621.002

; CURRENT FILING DATE: 1999-12-30

; NUMBER OF SEQ ID NOS: 90

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 74

; LENGTH: 4689

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence:

; OTHER INFORMATION: gp160.modsF162.gag.modsF2

; US-09-475-515-74

Query Match 97.3%; Score 58.4; DB 4; Length 4689;

Best Local Similarity 98.3%; Pred. No. 4e-10;

Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GACATCCGCCAGGGCCCCCAAGAGCCCTTCGCGACTAGTGACCGCTTCTTCAAGACC 60  
Db 4011 GACATCCGCCAGGGCCCCCAAGAGCCCTTCGCGACTAGTGACCGCTTCTTCAAGACC 4070

## RESULT 12

US-09-475-515-73

; Sequence 73, Application US/09475515A

; Patent No. 6602705

; GENERAL INFORMATION:

; APPLICANT: BARNETT, Susan

; APPLICANT: ZUR MEGEDE, Jan

; APPLICANT: SRIVASTAVA, Indresh

; APPLICANT: LIAN, Ying

; APPLICANT: HARTOG, Karin

; APPLICANT: LIU, Hong

; APPLICANT: GREER, Catherine

; APPLICANT: SELBY, Mark

; APPLICANT: WALKER, Christopher

; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION

; FILE REFERENCE: 1621.002

; CURRENT FILING DATE: 1999-12-30

; NUMBER OF SEQ ID NOS: 90

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 73

; LENGTH: 4766

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence:

; OTHER INFORMATION: gp160.modsU4.gag.modsF2

; US-09-475-515-73

; US-09-475-515-73

Query Match 97.3%; Score 58.4; DB 4; Length 4766;

Best Local Similarity 98.3%; Pred. No. 4e-10;

Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GACATCCGCCAGGGCCCCCAAGAGCCCTTCGCGACTAGTGACCGCTTCTTCAAGACC 60  
Db 4088 GACATCCGCCAGGGCCCCCAAGAGCCCTTCGCGACTAGTGACCGCTTCTTCAAGACC 4147

## RESULT 13

US-09-552-950-2

; Sequence 2, Application US/09552950

; Patent No. 6541248

; GENERAL INFORMATION:

; APPLICANT: Oxford Biomedica (UK) Limited

; APPLICANT: Oxford Biomedica (UK) Limited

; TITLE OF INVENTION: Anti-Viral Vectors

; FILE REFERENCE: 674524-2004

; CURRENT FILING DATE: 2000-04-20

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2

; LENGTH: 4307

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence:gagpol-SYNgp - codon

; US-09-552-950-2

Query Match 79.7%; Score 47.8; DB 4; Length 4307;



Best Local Similarity 88.1%; Pred. No. 9.9e-07;  
Matches 52; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 1 GACATCCGCCAGGCCCCCAAGAGCCCTTCGCACTAGTGACCGCTTCTCAAGAC 59  
DB 850 GACATCCGCCAGGCCCCCAAGAGACCTTTCGCACTAGTGACCGGTTCTACAAAC 908

## RESULT 14

US-09-552-950-5  
Sequence 5, Application US/09552950  
Patent No. 6541248

GENERAL INFORMATION:

APPLICANT: Oxford Biomedica (UK) Limited  
TITLE OF INVENTION: Anti-Viral Vectors

FILE REFERENCE: 674524-2004  
CURRENT APPLICATION NUMBER: US/09/552,950

CURRENT FILING DATE: 2000-04-20  
NUMBER OF SEQ ID NOS: 22

SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 5

LENGTH: 9772  
TYPE: DNA

ORGANISM: Artificial Sequence  
FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: pSYNGP  
US-09-552-950-5

Query Match 79.7%; Score 47.8; DB 4; Length 9772;  
Best Local Similarity 88.1%; Pred. No. 1.1e-06;  
Matches 52; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 1 GACATCCGCCAGGCCCCCAAGAGCCCTTCGCACTAGTGACCGCTTCTCAAGAC 59  
DB 1957 GACATCCGCCAGGCCCCCAAGAGACCTTTCGCACTAGTGACCGGTTCTACAAAC 2015

## RESULT 15

US-09-184-418C-100  
Sequence 100, Application US/09184418C  
Patent No. 6492110

GENERAL INFORMATION:

APPLICANT: Hahn, Beatrice  
APPLICANT: Gao, Feng

APPLICANT: Shaw, George  
TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN

FILE REFERENCE: D6287  
CURRENT APPLICATION NUMBER: US/09/184,418C

CURRENT FILING DATE: 1999-11-02  
NUMBER OF SEQ ID NOS: 112

SEQ ID NO 100  
LENGTH: 1476

TYPE: DNA  
ORGANISM: Human immunodeficiency virus type 1  
FEATURE:

OTHER INFORMATION: Isolate=94IN476.104; gene=gag  
US-09-184-418C-100

Query Match 58.3%; Score 35; DB 4; Length 1476;  
Best Local Similarity 74.6%; Pred. No. 0.011;  
Matches 44; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

OY 1 GACATCCGCCAGGCCCCCAAGAGCCCTTCGCACTAGTGACCGCTTCTCAAGAC 59  
DB 841 GACATAGACAGGCGCCAAAGAACCTTTAGAGACTATGTAGACCGGTTCTTAAAC 899

Search completed: May 28, 2004, 13:26:22  
Job time : 6.49156 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: May 28, 2004, 11:33:50 ; Search time 30.7692 Seconds  
(without alignments)  
8870.358 Million cell updates/sec

Title: US-09-475-704A-2  
Perfect score: 60  
Sequence: 1 gacatccgccaggcccaaa.....tgcacgccttctcaagacc 60

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2960401 seqs, 2274450654 residues

Total number of hits satisfying chosen parameters: 5920802

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq2:\*  
14: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*  
17: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
18: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
19: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	13 US-09-899-575-2	Sequence 2, Appli
2	60	100.0	1509	10 US-09-967-464-64	Sequence 64, Appli
3	60	100.0	1509	10 US-09-967-464-68	Sequence 68, Appli
4	60	100.0	1509	13 US-09-899-575-4	Sequence 4, Appli
5	60	100.0	1509	13 US-09-899-575-21	Sequence 21, Appli
6	60	100.0	1608	15 US-10-168-843A-27	Sequence 27, Appli
7	60	100.0	1914	15 US-10-168-843A-29	Sequence 29, Appli
8	60	100.0	2493	15 US-10-168-843A-31	Sequence 31, Appli
9	60	100.0	4350	15 US-10-168-843A-33	Sequence 33, Appli
10	58.4	97.3	60	13 US-10-387-336-20	Sequence 20, Appli
11	58.4	97.3	1268	13 US-10-387-336-9	Sequence 9, Appli
12	58.4	97.3	1515	13 US-10-387-336-4	Sequence 4, Appli
13	58.4	97.3	1853	13 US-10-387-336-5	Sequence 5, Appli
14	58.4	97.3	1865	13 US-10-387-336-78	Sequence 78, Appli

15	58.4	97.3	1865	13	US-10-387-336-79	Sequence 79, Appli
16	58.4	97.3	2031	13	US-10-387-336-7	Sequence 7, Appli
17	58.4	97.3	2799	15	US-10-241-009-18	Sequence 18, Appli
18	58.4	97.3	2799	15	US-10-241-009-53	Sequence 53, Appli
19	58.4	97.3	2799	15	US-10-190-434B-18	Sequence 18, Appli
20	58.4	97.3	2799	15	US-10-190-434B-53	Sequence 53, Appli
21	58.4	97.3	2799	15	US-10-190-434B-11	Sequence 11, Appli
22	58.4	97.3	2799	15	US-10-190-305A-74	Sequence 74, Appli
23	58.4	97.3	3205	15	US-10-241-009-17	Sequence 17, Appli
24	58.4	97.3	3205	15	US-10-190-434B-17	Sequence 17, Appli
25	58.4	97.3	3496	15	US-10-241-009-15	Sequence 15, Appli
26	58.4	97.3	3496	15	US-10-190-434B-15	Sequence 15, Appli
27	58.4	97.3	3564	15	US-10-241-009-13	Sequence 13, Appli
28	58.4	97.3	3564	15	US-10-241-009-14	Sequence 14, Appli
29	58.4	97.3	3564	15	US-10-190-434B-13	Sequence 13, Appli
30	58.4	97.3	3564	15	US-10-190-434B-14	Sequence 14, Appli
31	58.4	97.3	3999	15	US-10-241-009-9	Sequence 9, Appli
32	58.4	97.3	3999	15	US-10-241-009-10	Sequence 10, Appli
33	58.4	97.3	3999	15	US-10-241-009-11	Sequence 11, Appli
34	58.4	97.3	3999	15	US-10-190-434B-9	Sequence 9, Appli
35	58.4	97.3	3999	15	US-10-190-434B-10	Sequence 10, Appli
36	58.4	97.3	3999	15	US-10-190-434B-11	Sequence 11, Appli
37	58.4	97.3	4472	13	US-10-387-336-6	Sequence 6, Appli
38	58.4	97.3	4608	13	US-10-387-336-75	Sequence 75, Appli
39	58.4	97.3	4689	13	US-10-387-336-74	Sequence 74, Appli
40	58.4	97.3	4766	13	US-10-387-336-73	Sequence 73, Appli
41	58.4	97.3	4773	15	US-10-241-009-16	Sequence 16, Appli
42	58.4	97.3	4773	15	US-10-241-009-55	Sequence 55, Appli
43	58.4	97.3	4773	15	US-10-190-434B-16	Sequence 16, Appli
44	58.4	97.3	4773	15	US-10-190-434B-55	Sequence 55, Appli
45	58.4	97.3	4773	15	US-10-190-434B-55	Sequence 55, Appli

## ALIGNMENTS

RESULT 1  
US-09-899-575-2  
; Sequence 2, Application US/09899575  
; Publication No. US20030223961A1  
; GENERAL INFORMATION:  
; APPLICANT: Zur Megele, Jan  
; APPLICANT: Barnett, Susan W.  
; APPLICANT: Egnelbrecht, Susan  
; APPLICANT: van Renburg, Strelita Janse  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C  
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF  
; FILE REFERENCE: PP01631.102  
; CURRENT APPLICATION NUMBER: US/09/899,575  
; CURRENT FILING DATE: 2001-07-05  
; PRIOR APPLICATION NUMBER: 09/475,704  
; PRIOR FILING DATE: 1999-12-30  
; NUMBER OF SEQ ID NOS: 135  
; SOFTWARE: Patent Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 60  
; TYPE: DNA  
; ORGANISM: Human immunodeficiency virus  
US-09-899-575-2

Query Match 100.0%; Score 60; DB 13; Length 60;  
Best Local Similarity 100.0%; Pred. No. 2.2e-11;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATCCGCCAGGGCCCAAGAGCCCTTCGCGACTACGTGACCGCTTCAAGACC 60  
Db 1 GACATCCGCCAGGGCCCAAGAGCCCTTCGCGACTACGTGACCGCTTCAAGACC 60

RESULT 2  
US-09-967-464-64  
; Sequence 64, Application US/09967464  
; Publication No. US20030138453A1

GENERAL INFORMATION:  
APPLICANT: O'Hagan, Derek  
APPLICANT: Otten, Gillis  
APPLICANT: Donnelly, John J.  
APPLICANT: Polo, John M.  
APPLICANT: Barnett, Susan  
APPLICANT: Singh, Mamohan  
APPLICANT: Ulmer, Jeffrey  
APPLICANT: Dubensky, Jr., Thomas W.  
TITLE OF INVENTION: MICROPARTICLES FOR DELIVERY OF HETEROLOGOUS NUCLEIC ACIDS  
FILE REFERENCE: PP16269.004  
CURRENT APPLICATION NUMBER: US/09/967,464  
CURRENT FILING DATE: 2002-04-11  
PRIOR APPLICATION NUMBER: 60/236,105  
PRIOR FILING DATE: 2000-09-28  
PRIOR APPLICATION NUMBER: 60/315,905  
PRIOR FILING DATE: 2001-08-30  
NUMBER OF SEQ ID NOS: 68  
SOFTWARE: Patentin version 3.1  
SEQ ID NO 64  
LENGTH: 1509  
TYPE: DNA  
ORGANISM: Human immunodeficiency virus type 1  
US-09-967-464-64

Query Match 100.0%; Score 60; DB 10; Length 1509;  
Best Local Similarity 100.0%; Pred. No. 1.9e-11;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATCCGCCAGGCCCCCAAGAGAGCCCTTCCGCGACTACGTGAGCCGCTTCTTCAAGACC 60  
DB 841 GACATCCGCCAGGCCCCCAAGAGAGCCCTTCCGCGACTACGTGAGCCGCTTCTTCAAGACC 900

## RESULT 3

US-09-967-464-68  
Sequence 68, Application US/09967464  
Publication No. US20030138453A1

GENERAL INFORMATION:  
APPLICANT: O'Hagan, Derek  
APPLICANT: Otten, Gillis  
APPLICANT: Donnelly, John J.  
APPLICANT: Polo, John M.  
APPLICANT: Barnett, Susan  
APPLICANT: Singh, Mamohan  
APPLICANT: Ulmer, Jeffrey  
APPLICANT: Dubensky, Jr., Thomas W.  
TITLE OF INVENTION: MICROPARTICLES FOR DELIVERY OF HETEROLOGOUS NUCLEIC ACIDS  
FILE REFERENCE: PP16269.004  
CURRENT APPLICATION NUMBER: US/09/967,464  
CURRENT FILING DATE: 2002-04-11  
PRIOR APPLICATION NUMBER: 60/236,105  
PRIOR FILING DATE: 2000-09-28  
PRIOR APPLICATION NUMBER: 60/315,905  
PRIOR FILING DATE: 2001-08-30  
NUMBER OF SEQ ID NOS: 68  
SOFTWARE: Patentin version 3.1  
SEQ ID NO 68  
LENGTH: 1509  
TYPE: DNA  
ORGANISM: Human immunodeficiency virus type 1  
US-09-967-464-68

Query Match 100.0%; Score 60; DB 10; Length 1509;  
Best Local Similarity 100.0%; Pred. No. 1.9e-11;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATCCGCCAGGCCCCCAAGAGAGCCCTTCCGCGACTACGTGAGCCGCTTCTTCAAGACC 60  
DB 841 GACATCCGCCAGGCCCCCAAGAGAGCCCTTCCGCGACTACGTGAGCCGCTTCTTCAAGACC 900

## RESULT 4

US-09-899-575-4  
Sequence 4, Application US/09899575  
Publication No. US20030223961A1  
GENERAL INFORMATION:  
APPLICANT: Zur Megede, Jan  
APPLICANT: Barnett, Susan W.  
APPLICANT: Egnelbrecht, Susan  
APPLICANT: van Rensburg, Estrelita Janse  
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C  
TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF  
FILE REFERENCE: PP01631.102  
CURRENT APPLICATION NUMBER: US/09/899,575  
CURRENT FILING DATE: 2001-07-05  
PRIOR APPLICATION NUMBER: 09/475,704  
PRIOR FILING DATE: 1999-12-30  
NUMBER OF SEQ ID NOS: 135  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 4  
LENGTH: 1509  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURES:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic Gag  
US-09-899-575-4

Query Match 100.0%; Score 60; DB 13; Length 1509;  
Best Local Similarity 100.0%; Pred. No. 1.9e-11;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATCCGCCAGGCCCCCAAGAGAGCCCTTCCGCGACTACGTGAGCCGCTTCTTCAAGACC 60  
DB 841 GACATCCGCCAGGCCCCCAAGAGAGCCCTTCCGCGACTACGTGAGCCGCTTCTTCAAGACC 900

## RESULT 5

US-09-899-575-21  
Sequence 21, Application US/09899575  
Publication No. US20030223961A1

GENERAL INFORMATION:  
APPLICANT: Zur Megede, Jan  
APPLICANT: Barnett, Susan W.  
APPLICANT: Egnelbrecht, Susan  
APPLICANT: van Rensburg, Estrelita Janse  
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C  
TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF  
FILE REFERENCE: PP01631.102  
CURRENT APPLICATION NUMBER: US/09/899,575  
CURRENT FILING DATE: 2001-07-05  
PRIOR APPLICATION NUMBER: 09/475,704  
PRIOR FILING DATE: 1999-12-30  
NUMBER OF SEQ ID NOS: 135  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 21  
LENGTH: 1509  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURES:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic Gag  
US-09-899-575-21

Query Match 100.0%; Score 60; DB 13; Length 1509;  
Best Local Similarity 100.0%; Pred. No. 1.9e-11;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATCCGCCAGGCCCCCAAGAGAGCCCTTCCGCGACTACGTGAGCCGCTTCTTCAAGACC 60  
DB 841 GACATCCGCCAGGCCCCCAAGAGAGCCCTTCCGCGACTACGTGAGCCGCTTCTTCAAGACC 900

## RESULT 6

US-10-168-843A-27

```
; Sequence 27, Application US/10168843A
; Publication No. US20030108562A1
; GENERAL INFORMATION:
; APPLICANT: Medical Research Council
; APPLICANT: International Aids Vaccine Initiative
; APPLICANT: University of Nairobi
; TITLE OF INVENTION: Improvements in or Relating to Immune Responses to HIV
; FILE REFERENCE: MJL/C1248'1/M
; CURRENT APPLICATION NUMBER: US/10/168,843A
; CURRENT FILING DATE: 2002-09-24
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 1608
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Chimeric
; OTHER INFORMATION: polynucleotide
US-10-168-843A-27
```

```
Query Match          100.0%; Score 60; DB 15; Length 1608;
Best Local Similarity 100.0%; Pred. No. 1.9e-11;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 GACATCCGCCAGGGCCCCCAAGAGAGCCCTTCGCGACTAGCGACCGCTTCTTCAAGACC 60
        |||
Db      472 GACATCCGCCAGGGCCCCCAAGAGAGCCCTTCGCGACTAGCGACCGCTTCTTCAAGACC 531
```

## RESULT 7

```
US-10-168-843A-29
; Sequence 29, Application US/10168843A
; Publication No. US20030108562A1
; GENERAL INFORMATION:
; APPLICANT: Medical Research Council
; APPLICANT: International Aids Vaccine Initiative
; APPLICANT: University of Nairobi
; TITLE OF INVENTION: Improvements in or Relating to Immune Responses to HIV
; FILE REFERENCE: MJL/C1248'1/M
; CURRENT APPLICATION NUMBER: US/10/168,843A
; CURRENT FILING DATE: 2002-09-24
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29
; LENGTH: 1914
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Chimeric
; OTHER INFORMATION: polynucleotide
US-10-168-843A-29
```

```
Query Match          100.0%; Score 60; DB 15; Length 1914;
Best Local Similarity 100.0%; Pred. No. 1.9e-11;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 GACATCCGCCAGGGCCCCCAAGAGAGCCCTTCGCGACTAGCGACCGCTTCTTCAAGACC 60
        |||
Db      466 GACATCCGCCAGGGCCCCCAAGAGAGCCCTTCGCGACTAGCGACCGCTTCTTCAAGACC 525
```

## RESULT 8

```
US-10-168-843A-31
; Sequence 31, Application US/10168843A
; Publication No. US20030108562A1
; GENERAL INFORMATION:
; APPLICANT: Medical Research Council
; APPLICANT: International Aids Vaccine Initiative
; APPLICANT: University of Nairobi
; TITLE OF INVENTION: Improvements in or Relating to Immune Responses to HIV
; FILE REFERENCE: MJL/C1248'1/M
; CURRENT APPLICATION NUMBER: US/10/168,843A
```

```
; CURRENT FILING DATE: 2002-09-24
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 31
; LENGTH: 2493
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Chimeric
; OTHER INFORMATION: polynucleotide
US-10-168-843A-31
```

```
Query Match          100.0%; Score 60; DB 15; Length 2493;
Best Local Similarity 100.0%; Pred. No. 1.9e-11;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 GACATCCGCCAGGGCCCCCAAGAGAGCCCTTCGCGACTAGCGACCGCTTCTTCAAGACC 60
        |||
Db      466 GACATCCGCCAGGGCCCCCAAGAGAGCCCTTCGCGACTAGCGACCGCTTCTTCAAGACC 525
```

## RESULT 9

```
US-10-168-843A-33
; Sequence 33, Application US/10168843A
; Publication No. US20030108562A1
; GENERAL INFORMATION:
; APPLICANT: Medical Research Council
; APPLICANT: International Aids Vaccine Initiative
; APPLICANT: University of Nairobi
; TITLE OF INVENTION: Improvements in or Relating to Immune Responses to HIV
; FILE REFERENCE: MJL/C1248'1/M
; CURRENT APPLICATION NUMBER: US/10/168,843A
; CURRENT FILING DATE: 2002-09-24
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 33
; LENGTH: 4350
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Chimeric
; OTHER INFORMATION: polynucleotide
US-10-168-843A-33
```

```
Query Match          100.0%; Score 60; DB 15; Length 4350;
Best Local Similarity 100.0%; Pred. No. 1.8e-11;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 GACATCCGCCAGGGCCCCCAAGAGAGCCCTTCGCGACTAGCGACCGCTTCTTCAAGACC 60
        |||
Db      466 GACATCCGCCAGGGCCCCCAAGAGAGCCCTTCGCGACTAGCGACCGCTTCTTCAAGACC 525
```

## RESULT 10

```
US-10-387-336-20
; Sequence 20, Application US/10387336
; Publication No. US20030223964A1
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/10/387,336
; CURRENT FILING DATE: 2003-03-11
; PRIOR APPLICATION NUMBER: US/09/475,515A
```

PRIOR FILING DATE: 1999-12-30  
NUMBER OF SEQ ID NOS: 90  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 20  
LENGTH: 60  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic p55  
US-10-387-336-20

Query Match 97.3%; Score 58.4; DB 13; Length 60;  
Best Local Similarity 98.3%; Pred. No. 7.7e-11;  
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GACATCCGCGCAGGGCCCCCAAGAGACCCCTTCGCGACTACGTGACCGCTTCTTCAAGACC 60  
DB 1 GACATCCGCGCAGGGCCCCCAAGAGACCCCTTCGCGACTACGTGACCGCTTCTTCAAGACC 60

RESULT 11  
US-10-387-336-9  
Sequence 9, Application US/10387336  
Publication No. US20030223964A1  
GENERAL INFORMATION:  
APPLICANT: BARNETT, Susan  
APPLICANT: ZUR MEGEDE, Jan  
APPLICANT: SRIVASTAVA, Indresh  
APPLICANT: LIAN, Ying  
APPLICANT: HARTOG, Karin  
APPLICANT: LIU, Hong  
APPLICANT: GREER, Catherine  
APPLICANT: SELBY, Mark  
APPLICANT: WALKER, Christopher  
TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION  
TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES  
FILE REFERENCE: 1621.002  
CURRENT APPLICATION NUMBER: US/10/387,336  
CURRENT FILING DATE: 2003-03-11  
PRIOR APPLICATION NUMBER: US/09/475,515A  
PRIOR FILING DATE: 1999-12-30  
NUMBER OF SEQ ID NOS: 90  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 9  
LENGTH: 1268  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic Gag  
US-10-387-336-9

Query Match 97.3%; Score 58.4; DB 13; Length 1268;  
Best Local Similarity 98.3%; Pred. No. 6.8e-11;  
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GACATCCGCGCAGGGCCCCCAAGAGACCCCTTCGCGACTACGTGACCGCTTCTTCAAGACC 60  
DB 862 GACATCCGCGCAGGGCCCCCAAGAGACCCCTTCGCGACTACGTGACCGCTTCTTCAAGACC 921

RESULT 12  
US-10-387-336-4  
Sequence 4, Application US/10387336  
Publication No. US20030223964A1  
GENERAL INFORMATION:  
APPLICANT: BARNETT, Susan  
APPLICANT: ZUR MEGEDE, Jan  
APPLICANT: SRIVASTAVA, Indresh  
APPLICANT: LIAN, Ying  
APPLICANT: HARTOG, Karin  
APPLICANT: LIU, Hong

APPLICANT: GREER, Catherine  
APPLICANT: SELBY, Mark  
APPLICANT: WALKER, Christopher  
TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION  
TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES  
FILE REFERENCE: 1621.002  
CURRENT APPLICATION NUMBER: US/10/387,336  
CURRENT FILING DATE: 2003-03-11  
PRIOR APPLICATION NUMBER: US/09/475,515A  
PRIOR FILING DATE: 1999-12-30  
NUMBER OF SEQ ID NOS: 90  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 4  
LENGTH: 1515  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-10-387-336-4

Query Match 97.3%; Score 58.4; DB 13; Length 1515;  
Best Local Similarity 98.3%; Pred. No. 6.8e-11;  
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GACATCCGCGCAGGGCCCCCAAGAGACCCCTTCGCGACTACGTGACCGCTTCTTCAAGACC 60  
DB 862 GACATCCGCGCAGGGCCCCCAAGAGACCCCTTCGCGACTACGTGACCGCTTCTTCAAGACC 921

RESULT 13  
US-10-387-336-5  
Sequence 5, Application US/10387336  
Publication No. US20030223964A1  
GENERAL INFORMATION:  
APPLICANT: BARNETT, Susan  
APPLICANT: ZUR MEGEDE, Jan  
APPLICANT: SRIVASTAVA, Indresh  
APPLICANT: LIAN, Ying  
APPLICANT: HARTOG, Karin  
APPLICANT: LIU, Hong  
APPLICANT: GREER, Catherine  
APPLICANT: SELBY, Mark  
APPLICANT: WALKER, Christopher  
TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION  
TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES  
FILE REFERENCE: 1621.002  
CURRENT APPLICATION NUMBER: US/10/387,336  
CURRENT FILING DATE: 2003-03-11  
PRIOR APPLICATION NUMBER: US/09/475,515A  
PRIOR FILING DATE: 1999-12-30  
NUMBER OF SEQ ID NOS: 90  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 5  
LENGTH: 1853  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-10-387-336-5

Query Match 97.3%; Score 58.4; DB 13; Length 1853;  
Best Local Similarity 98.3%; Pred. No. 6.7e-11;  
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GACATCCGCGCAGGGCCCCCAAGAGACCCCTTCGCGACTACGTGACCGCTTCTTCAAGACC 60  
DB 862 GACATCCGCGCAGGGCCCCCAAGAGACCCCTTCGCGACTACGTGACCGCTTCTTCAAGACC 921

RESULT 14  
US-10-387-336-78



; Sequence 78, Application US/10387336  
; Publication No. US20030223964A1  
; GENERAL INFORMATION:  
; APPLICANT: BARNETT, Susan  
; APPLICANT: ZUR MEGEDE, Jan  
; APPLICANT: SRIVASTAVA, Indresh  
; APPLICANT: LIAN, Ying  
; APPLICANT: HARTOG, Karin  
; APPLICANT: LIU, Hong  
; APPLICANT: GREER, Catherine  
; APPLICANT: SELBY, Mark  
; APPLICANT: WALKER, Christopher  
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION  
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES  
; FILE REFERENCE: 1621.002  
; CURRENT APPLICATION NUMBER: US/10/387,336  
; CURRENT FILING DATE: 2003-03-11  
; PRIOR APPLICATION NUMBER: US/09/475,515A  
; PRIOR FILING DATE: 1999-12-30  
; NUMBER OF SEQ ID NOS: 90  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 78  
; LENGTH: 1865  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: GP1  
US-10-387-336-78

Query Match 97.3%; Score 58.4; DB 13; Length 1865;  
Best Local Similarity 98.3%; Pred. No. 6.7e-11;  
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GACATCCGCCAGGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 60  
|||||  
DB 868 GACATCCGCCAGGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 927

RESULT 15  
US-10-387-336-79  
; Sequence 79, Application US/10387336  
; Publication No. US20030223964A1  
; GENERAL INFORMATION:  
; APPLICANT: BARNETT, Susan  
; APPLICANT: ZUR MEGEDE, Jan  
; APPLICANT: SRIVASTAVA, Indresh  
; APPLICANT: LIAN, Ying  
; APPLICANT: HARTOG, Karin  
; APPLICANT: LIU, Hong  
; APPLICANT: GREER, Catherine  
; APPLICANT: SELBY, Mark  
; APPLICANT: WALKER, Christopher  
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION  
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES  
; FILE REFERENCE: 1621.002  
; CURRENT APPLICATION NUMBER: US/10/387,336  
; CURRENT FILING DATE: 2003-03-11  
; PRIOR APPLICATION NUMBER: US/09/475,515A  
; PRIOR FILING DATE: 1999-12-30  
; NUMBER OF SEQ ID NOS: 90  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 79  
; LENGTH: 1865  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: GP2  
US-10-387-336-79

Query Match 97.3%; Score 58.4; DB 13; Length 1865;  
Best Local Similarity 98.3%; Pred. No. 6.7e-11;  
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GACATCCGCCAGGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 60  
|||||  
DB 868 GACATCCGCCAGGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 927

Search completed: May 28, 2004, 16:09:15  
Job time : 30.7692 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 28, 2004, 07:10:24 ; Search time 27.8049 Seconds  
(without alignments)  
9167.164 Million cell updates/sec

Title: US-09-475-704A-2

Perfect score: 60  
Sequence: 1 gacatccgcagggcccccacaa.....tggaccgcttctcaagacc 60

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_29Jan04:\*

1:	geneseqn1980s:*
2:	geneseqn1990s:*
3:	geneseqn2000s:*
4:	geneseqn2001as:*
5:	geneseqn2001bs:*
6:	geneseqn2002as:*
7:	geneseqn2003as:*
8:	geneseqn2003bs:*
9:	geneseqn2003cs:*
10:	geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	3	AA51608 HIV Gag m
2	60	100.0	1509	3	AA51610 HIV syth
3	60	100.0	1509	3	AA51626 HIV codon
4	60	100.0	1509	6	AA144553 HIV-1 p55
5	60	100.0	1509	6	AA144549 HIV-1 p55
6	60	100.0	1509	6	ABL39954 Synthetic
7	60	100.0	1509	6	ABL39958 Synthetic
8	60	100.0	1608	4	AA09487 Human imm
9	60	100.0	1914	4	AA09488 Human imm
10	60	100.0	2493	4	AA09489 Human imm
11	60	100.0	4350	4	AA09490 Human imm
12	58.4	97.3	60	3	AA70426 HIV p55 G
13	58.4	97.3	1268	3	AA70417 HIV Gag c
14	58.4	97.3	1503	3	AA252051 Codon opt
15	58.4	97.3	1515	3	AA70412 Synthetic
16	58.4	97.3	1853	3	AA70413 HIV Gag-p
17	58.4	97.3	1865	3	AA70475 HIV Gag-p
18	58.4	97.3	1865	3	AA70476 HIV Gag-p
19	58.4	97.3	2031	3	AA70415 Synthetic
20	58.4	97.3	2799	7	ACA03583 Synthetic
21	58.4	97.3	2799	7	ACA03520 Synthetic
22	58.4	97.3	2799	7	ACC78493 HIV GagTa
23	58.4	97.3	2799	7	ACC78528 HIV TatRe

24	58.4	97.3	3205	7	ACC78492	ACC78492 HIV GagRT
25	58.4	97.3	3496	7	ACC78490	ACC78490 HIV GagPr
26	58.4	97.3	3564	7	ACC78488	ACC78488 HIV GagPo
27	58.4	97.3	3564	7	ACC78489	ACC78489 HIV GagPo
28	58.4	97.3	3999	7	ACC78484	ACC78484 HIV GagCo
29	58.4	97.3	3999	7	ACC78485	ACC78485 HIV GagCo
30	58.4	97.3	3999	7	ACC78486	ACC78486 HIV GagCo
31	58.4	97.3	4319	3	AA70414	AA70414 HIV Gag-P
32	58.4	97.3	4472	3	AA70472	AA70472 HIV bicis
33	58.4	97.3	4608	3	AA70473	AA70473 HIV bicis
34	58.4	97.3	4689	3	AA70471	AA70471 HIV bicis
35	58.4	97.3	4766	3	AA70470	AA70470 HIV bicis
36	58.4	97.3	4773	7	ACA03519	ACA03519 Synthetic
37	58.4	97.3	4773	7	ACA03585	ACA03585 Synthetic
38	58.4	97.3	4773	7	ACC78491	ACC78491 HIV GagPr
39	58.4	97.3	4773	7	ACC78530	ACC78530 HIV TatRe
40	58.4	97.3	5274	7	ACA03518	ACA03518 Synthetic
41	58.4	97.3	5274	7	ACC78487	ACC78487 HIV GagCp
42	58.4	97.3	5283	7	ACA03584	ACA03584 Synthetic
43	58.4	97.3	5283	7	ACC78529	ACC78529 HIV TatRe
44	58.4	97.3	6438	6	ABK91605	ABK91605 Modified
45	58.4	97.3	8908	3	AA252055	AA252055 Packaging

## ALIGNMENTS

RESULT 1  
AAA51608  
ID AAA51608 standard; DNA; 60 BP.  
XX  
AC AAA51608;  
XX  
DT 31-OCT-2000 (first entry)  
XX  
DE HIV Gag major homology region nucleotides 841-900.  
XX  
KW Gag; expression cassette; antigenic; type C; HIV; Env; synthetic;  
KW DNA immunization; packaging cell line; antigen presentation; ss.  
XX  
OS Human immunodeficiency virus.  
XX  
PN WO200039304-A2.  
XX  
PD 06-JUL-2000.  
XX  
PF 30-DEC-1999; 99WO-US031273.  
XX  
PR 31-DEC-1998; 98US-0114495P.  
PR 01-SEP-1999; 99US-0152195P.  
XX  
PA (CHIR ) CHIRON CORP.  
XX  
PI Barnett S, Zur Megede J;  
XX  
DR WPI; 2000-452401/39.  
XX  
PT Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV Env  
PT polypeptide and the polypeptide useful for immunizing a mammal especially  
PT human against HIV.  
XX  
PS Claim 1; Page 92; 113pp; English.  
XX  
CC Expression cassettes comprising a polynucleotide encoding antigenic type  
CC C human immunodeficiency virus (HIV) Gag or Env polypeptides are useful  
CC in DNA immunization, generation of packaging cell lines and production of  
CC Gag- and/or Env-containing proteins. Synthetic Env and Gag expression  
CC cassettes exhibit increased potency for induction of cytotoxic T-  
CC lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1 self-  
CC assemble into non-infectious virus-like particles which are used as a  
CC matrix for the proper presentation of an antigen entrapped or associated  
CC to the immune system of the host

8Q Sequence 60 BP; 12 A; 24 C; 15 G; 9 T; 0 U; 0 Other;

Query Match 100.0%; Score 60; DB 3; Length 60;  
Best Local Similarity 100.0%; Pred. No. 4.5e-09;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GACATCCGCCAGGCCCCCAAGAGACCCCTTCGCGACTACGTGACCGCTTCTTCAAGACC 60  
1 GACATCCGCCAGGCCCCCAAGAGACCCCTTCGCGACTACGTGACCGCTTCTTCAAGACC 60  
DB

## RESULT 2

AAA51610 standard; DNA; 1509 BP.

AC AAA51610;

DT 31-OCT-2000 (first entry)

DE HIV synthetic Gag polynucleotide.

KW Gag; expression cassette; antigenic; type C; HIV; Env; synthetic;  
KM DNA immunization; packaging cell line; antigen presentation; ss.

OS Human immunodeficiency virus; type C strain AF110967.  
Synthetic.

PN WO200039304-A2.

PD 06-JUL-2000.

PF 30-DEC-1999; 99WO-US031273.

PR 31-DEC-1998; 98US-0114495P.  
01-SEP-1998; 99US-0152195P.

PA (CHIR ) CHIRON CORP.

PI Barnett S, Zur Megede J;

WPI; 2000-452401/39.

PT Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV Env  
PT polypeptide and the polypeptide useful for immunizing a mammal especially  
PT human against HIV.

PS Claim 2; Page 93; 113pp; English.

CC Expression cassettes comprising a polynucleotide encoding antigenic type  
CC C human immunodeficiency virus (HIV) Gag or Env polypeptides are useful  
CC in DNA immunization, generation of packaging cell lines and production of  
CC Gag- and/or Env-containing proteins. Synthetic Env and Gag expression  
CC cassettes exhibit increased potency for induction of cytotoxic T-  
CC lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1 self-  
CC assemble into non-infectious virus-like particles which are used as a  
CC matrix for the proper presentation of an antigen entrapped or associated  
CC to the immune system of the host

8Q Sequence 1509 BP; 320 A; 556 C; 472 G; 161 T; 0 U; 0 Other;

Query Match 100.0%; Score 60; DB 3; Length 1509;  
Best Local Similarity 100.0%; Pred. No. 5.7e-09;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GACATCCGCCAGGCCCCCAAGAGACCCCTTCGCGACTACGTGACCGCTTCTTCAAGACC 60  
1 GACATCCGCCAGGCCCCCAAGAGACCCCTTCGCGACTACGTGACCGCTTCTTCAAGACC 60  
DB 841 GACATCCGCCAGGCCCCCAAGAGACCCCTTCGCGACTACGTGACCGCTTCTTCAAGACC 900

## RESULT 3

AAA51626

ID AAA51626 standard; DNA; 1509 BP.  
XX

AC AAA51626;

DT 31-OCT-2000 (first entry)

DB HIV codon-optimized synthetic Gag polynucleotide.

KW Gag; expression cassette; antigenic; type C; HIV; Env; synthetic;  
KM DNA immunization; packaging cell line; antigen presentation; ss.

OS Human immunodeficiency virus; type C strain AF110967.  
Synthetic.

PN WO200039304-A2.

PD 06-JUL-2000.

PF 30-DEC-1999; 99WO-US031273.

PR 31-DEC-1998; 98US-0114495P.  
01-SEP-1999; 99US-0152195P.

PA (CHIR ) CHIRON CORP.

PI Barnett S, Zur Megede J;

WPI; 2000-452401/39.

PT Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV Env  
PT polypeptide and the polypeptide useful for immunizing a mammal especially  
PT human against HIV.

PS Disclosure; Page 104; 113pp; English.

CC Expression cassettes comprising a polynucleotide encoding antigenic type  
CC C human immunodeficiency virus (HIV) Gag or Env polypeptides are useful  
CC in DNA immunization, generation of packaging cell lines and production of  
CC Gag- and/or Env-containing proteins. Synthetic Env and Gag expression  
CC cassettes exhibit increased potency for induction of cytotoxic T-  
CC lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1 self-  
CC assemble into non-infectious virus-like particles which are used as a  
CC matrix for the proper presentation of an antigen entrapped or associated  
CC to the immune system of the host

8Q Sequence 1509 BP; 321 A; 559 C; 471 G; 158 T; 0 U; 0 Other;

Query Match 100.0%; Score 60; DB 3; Length 1509;  
Best Local Similarity 100.0%; Pred. No. 5.7e-09;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GACATCCGCCAGGCCCCCAAGAGACCCCTTCGCGACTACGTGACCGCTTCTTCAAGACC 60  
1 GACATCCGCCAGGCCCCCAAGAGACCCCTTCGCGACTACGTGACCGCTTCTTCAAGACC 60  
DB 841 GACATCCGCCAGGCCCCCAAGAGACCCCTTCGCGACTACGTGACCGCTTCTTCAAGACC 900

## RESULT 4

AAL44553 standard; DNA; 1509 BP.

AC AAL44553;

DT 29-AUG-2003 (revised)  
DT 08-NOV-2002 (first entry)

DE HIV-1 p55gag polypeptide coding sequence 4.

KW HIV; ds; vaccine; gene; immune response; microparticle;  
KM adsorbent surface; poly(alpha-hydroxy acid); polyhydroxy butyric acid;  
KM polycaprolactone; polyorthoester; polycyanacrylate; detergent;  
KM submicron emulsion; viral infection; bacterial infection;  
KM parasitic infection; HIV-1 p55gag polypeptide.

OS Human immunodeficiency virus 1.

XX

```

FH Key Location/Qualifiers
FT variation replace(282, C)
FT variation /*tag= a
FT variation replace(331, T)
FT variation /*tag= b
FT variation replace(332, C)
FT variation /*tag= c
FT variation replace(549, G)
FT variation /*tag= d
FT variation replace(553, T)
FT variation /*tag= e
FT variation replace(783, G)
FT variation /*tag= f
FT variation replace(816, G)
FT variation /*tag= g
FT variation replace(999, T)
FT variation /*tag= h
FT variation replace(1002, C)
FT variation /*tag= i
FT variation replace(1089, G)
FT variation /*tag= j
FT variation replace(1149, G)
FT variation /*tag= k
FT variation replace(1158, C)
FT variation /*tag= l

XX WO200226209-A2.
XX
XX PD 04-APR-2002.
XX PF 28-SEP-2001; 2001WO-US030540.
XX PR 28-SEP-2000; 2000US-0236105P.
XX PR 30-AUG-2001; 2001US-0315905P.
XX PA (CHIR ) CHIRON CORP.
XX PI O'hagan D, Otten G, Donnelly JJ, Polo JM, Barnett S, Singh M;
XX PI Ulmer J, Dubensky TW;
XX DR WPI; 2002-519084/55.
XX
XX PT A microparticle to which a biologically active macromolecule is adsorbed,
XX PT for use as a vaccine composition to treat viral, bacterial or parasitic
XX PT infections, comprises a polymer microparticle, a detergent and a
XX PT submicron emulsion.
XX PS Claim 72; Fig 6; 100pp; English.
XX
CC The invention relates to a method of raising an immune response in a host
CC animal. The method of the invention comprises administering a
CC microparticle that has an adsorbent surface to which a first biologically
CC active macromolecule (e.g. a nucleic acid) has been adsorbed. The
CC microparticle comprises a polymer microparticle of poly(alpha-hydroxy
CC acid), a polyhydroxy butyric acid, a polycaprolactone, a polyorthoester,
CC a polycyanoacrylate, a detergent, and submicron emulsion. The method/
CC microparticle of the invention is useful for immunising a host animal
CC against viral, bacterial or parasitic infections. The present DNA
CC sequence encodes a HIV-1 p55gag polypeptide. (Updated on 29-AUG-2003 to
CC standardise OS field)
XX
SQ Sequence 1509 BP; 321 A; 559 C; 471 G; 158 T; 0 U; 0 Other;

Query Match 100.0%; Score 60; DB 6; Length 1509;
Best Local Similarity 100.0%; Pred. No. 5.7e-09;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GACATCCGCCAGGGCCCCCAAGAGGCCCTTCCGCGACTAGGTGACCGCTTCTTCAAGACC 60
DB 841 GACATCCGCCAGGGCCCCCAAGAGGCCCTTCCGCGACTAGGTGACCGCTTCTTCAAGACC 900

RESULT 5
```

```

AAL44549
ID AAL44549 standard; DNA; 1509 BP.
XX
XX AC AAL44549;
XX
XX DT 29-AUG-2003 (revised)
XX DT 08-NOV-2002 (first entry)
XX
XX DE HIV-1 p55gag polypeptide coding sequence 2.
XX
XX KW HIV; ds; vaccine; gene; immune response; microparticle;
XX KW adsorbent surface; poly(alpha-hydroxy acid); polyhydroxy butyric acid;
XX KW polycaprolactone; polyorthoester; polycyanoacrylate; detergent;
XX KW submicron emulsion; viral infection; bacterial infection;
XX KW parasitic infection; HIV-1 p55gag polypeptide.
XX
XX OS Human immunodeficiency virus 1.
XX
XX PN WO200226209-A2.
XX
XX PD 04-APR-2002.
XX
XX PF 28-SEP-2001; 2001WO-US030540.
XX
XX PR 28-SEP-2000; 2000US-0236105P.
XX PR 30-AUG-2001; 2001US-0315905P.
XX PA (CHIR ) CHIRON CORP.
XX
XX PI O'hagan D, Otten G, Donnelly JJ, Polo JM, Barnett S, Singh M;
XX PI Ulmer J, Dubensky TW;
XX DR WPI; 2002-519084/55.
XX
XX PT A microparticle to which a biologically active macromolecule is adsorbed,
XX PT for use as a vaccine composition to treat viral, bacterial or parasitic
XX PT infections, comprises a polymer microparticle, a detergent and a
XX PT submicron emulsion.
XX PS Claim 72; Fig 2; 100pp; English.
XX
CC The invention relates to a method of raising an immune response in a host
CC animal. The method of the invention comprises administering a
CC microparticle that has an adsorbent surface to which a first biologically
CC active macromolecule (e.g. a nucleic acid) has been adsorbed. The
CC microparticle comprises a polymer microparticle of poly(alpha-hydroxy
CC acid), a polyhydroxy butyric acid, a polycaprolactone, a polyorthoester,
CC a polycyanoacrylate, a detergent, and submicron emulsion. The method/
CC microparticle of the invention is useful for immunising a host animal
CC against viral, bacterial or parasitic infections. The present DNA
CC sequence encodes a HIV-1 p55gag polypeptide. (Updated on 29-AUG-2003 to
CC standardise OS field)
XX
SQ Sequence 1509 BP; 320 A; 556 C; 472 G; 161 T; 0 U; 0 Other;

Query Match 100.0%; Score 60; DB 6; Length 1509;
Best Local Similarity 100.0%; Pred. No. 5.7e-09;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GACATCCGCCAGGGCCCCCAAGAGGCCCTTCCGCGACTAGGTGACCGCTTCTTCAAGACC 60
DB 841 GACATCCGCCAGGGCCCCCAAGAGGCCCTTCCGCGACTAGGTGACCGCTTCTTCAAGACC 900

RESULT 6
ABL39954
ID ABL39954 standard; DNA; 1509 BP.
XX
XX AC ABL39954;
XX
XX DT 15-MAY-2002 (first entry)
XX
XX DT Synthetic Gag polynucleotide sequence SEQ ID NO:4.
XX
```

```
XX Human immunodeficiency virus type C; antigenic HIV type C protein;
KW immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
KW immunostimulant; gene therapy; gene; ds.
XX Human immunodeficiency virus; type C.
OS Synthetic.
XX WO200204493-A2.
XX 17-JAN-2002.
XX 05-JUL-2001; 2001WO-US021241.
XX 05-JUL-2000; 2000US-00610313.
XX (CHIR ) CHIRON CORP.
XX (UYST-) UNIV STELLENBOSCH.
XX Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg BJ;
XX WPI; 2002-154920/20.
XX WPI; 2002-154920/20.
XX New polynucleotides encoding antigenic HIV Type C polypeptides, useful in
XX applications including DNA immunization or generation of packaging cell
XX lines, particularly in gene therapy.
XX Example 1; Fig 2; 233pp; English.
XX The present invention describes expression cassettes comprising a
XX polynucleotide sequence encoding a polypeptide comprising immunogenic HIV
XX type C polypeptides. The expression cassettes comprise any of the HIV
XX type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or Nef
XX (I). (I) have immunostimulant activity and can be used in gene therapy.
XX The HIV type C polynucleotides are useful in applications including DNA
XX immunisation, generation of packaging cell lines, and production of HIV
XX type C proteins. The polynucleotides are particularly useful in gene
XX therapy and DNA immunisation applications. ABL39942 to ABL40054 and
XX ABB06204 to ABB06215 represent sequences used in the exemplification of
XX the present invention
XX Sequence 1509 BP; 320 A; 556 C; 472 G; 161 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 60; DB 6; Length 1509;
Best Local Similarity 100.0%; Pred. No. 5.7e-09;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GACATCCGCCAGGCCCCCAAGAGACCTTCGCGACTACGTGACCGCTTCTCAAGACC 60
DB 841 GACATCCGCCAGGCCCCCAAGAGACCTTCGCGACTACGTGACCGCTTCTCAAGACC 900
RESULT 7
ABL39958
ID ABL39958 standard; DNA; 1509 BP.
XX ABL39958;
XX AC
XX 15-MAY-2002 (first entry)
XX DT
XX Synthetic Gag polynucleotide sequence SEQ ID NO:21.
XX DB
XX Human immunodeficiency virus type C; antigenic HIV type C protein;
KW immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
KW immunostimulant; gene therapy; gene; ds.
XX Human immunodeficiency virus; type C.
OS Synthetic.
XX WO200204493-A2.
XX 17-JAN-2002.
XX
```

```
PF 05-JUL-2001; 2001WO-US021241.
XX 05-JUL-2000; 2000US-00610313.
XX (CHIR ) CHIRON CORP.
XX (UYST-) UNIV STELLENBOSCH.
XX Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg BJ;
XX WPI; 2002-154920/20.
XX New polynucleotides encoding antigenic HIV Type C polypeptides, useful in
XX applications including DNA immunization or generation of packaging cell
XX lines, particularly in gene therapy.
XX Example 1; Fig 6; 233pp; English.
XX The present invention describes expression cassettes comprising a
XX polynucleotide sequence encoding a polypeptide comprising immunogenic HIV
XX type C polypeptides. The expression cassettes comprise any of the HIV
XX type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or Nef
XX (I). (I) have immunostimulant activity and can be used in gene therapy.
XX The HIV type C polynucleotides are useful in applications including DNA
XX immunisation, generation of packaging cell lines, and production of HIV
XX type C proteins. The polynucleotides are particularly useful in gene
XX therapy and DNA immunisation applications. ABL39942 to ABL40054 and
XX ABB06204 to ABB06215 represent sequences used in the exemplification of
XX the present invention
XX Sequence 1509 BP; 321 A; 560 C; 470 G; 158 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 60; DB 6; Length 1509;
Best Local Similarity 100.0%; Pred. No. 5.7e-09;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GACATCCGCCAGGCCCCCAAGAGACCTTCGCGACTACGTGACCGCTTCTCAAGACC 60
DB 841 GACATCCGCCAGGCCCCCAAGAGACCTTCGCGACTACGTGACCGCTTCTCAAGACC 900
RESULT 8
AAD09487
ID AAD09487 standard; DNA; 1608 BP.
XX AAD09487;
XX AC
XX 10-SEP-2001 (first entry)
XX DT
XX Human immunodeficiency virus A (HIV A) gene.
XX DB
XX Human immunodeficiency virus A; HIV A; immunogen; anti-HIV; vaccine;
KW gene therapy; fusion protein; modified vaccinia virus Ankara vector; MVA;
KW cytotoxic T-lymphocyte; CTL; epitope; ds.
XX Human immunodeficiency virus.
OS
XX Key Location/Qualifiers
XX CDS 19..1602
XX FT /*tag= a
XX FT /product= "HIV A immunogen"
XX PN
XX WO200147955-A2.
XX 05-JUL-2001.
XX 22-DEC-2000; 2000WO-GB004984.
XX 23-DEC-1999; 99GB-00030294.
XX 14-OCT-2000; 2000GB-00025234.
XX (MEDI-) MEDICAL RES COUNCIL.
XX (ITAI-) INT AIDS VACCINE INITIATIVE.
XX (UYNA-) UNIV NAIROBI.
XX
```



```
XX PI Hanke T, Mcmichael AJ;
XX DR WPI; 2001-418221/44.
XX DR P-PSDB; AAE04825.
XX PT Novel immunogen for stimulating anti-HIV immune response, has a portion
XX PT of gag protein of HIV from HIV clade, parts of p17, p24 and synthetic
XX PT polypeptide comprising human cytotoxic T-lymphocyte epitopes of HIV
XX PT protein.
XX PS Claim 29; Fig 2A; 65pp; English.
XX CC The invention relates to human immunodeficiency virus immunogens and
XX CC their corresponding DNA molecules. An immunogen comprises a portion of
XX CC gag protein of HIV from an HIV clade, parts of p17 and p24, modified to
XX CC prevent N-terminal myristoylation; and a synthetic polypeptide comprising
XX CC human cytotoxic T-lymphocyte (CTL) epitopes of HIV protein. This
XX CC immunogen is designed to elicit an HIV-specific immune response in
XX CC humans. The immunogen is useful in the preparation of a medicament such
XX CC as vaccine to prevent or treat HIV infection in a human subject. The
XX CC invention also relates to method of stimulating anti-HIV immune response
XX CC in a human subject which comprises administering one or more times an
XX CC amount of nucleic acid molecule sufficient to prime an immune response to
XX CC the immunogen, or else may be packaged within a delivery means, such as a
XX CC modified vaccinia virus Ankara (MVA) to boost the immune response to
XX CC common portion of the immunogens. The present DNA sequence encodes human
XX CC immunodeficiency virus A immunogen (HIV A) fusion protein construct. HIV
XX CC A immunogen consists of about 73% of gag protein fused to a string of 25
XX CC partially overlapping human CTL epitopes. The gag domain of HIV A
XX CC contains p24 and p17 in an order reversed to the viral gag p17-p24-p15
XX CC polypeptide.
XX SQ Sequence 1608 BP; 314 A; 598 C; 458 G; 238 T; 0 U; 0 Other;
OY Query Match 100.0%; Score 60; DB 4; Length 1608;
Best Local Similarity 100.0%; Pred. No. 5.7e-09;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1 GACATCCGCCAGGGCCCAAGAGAGCCCTTCGGGACTGAGTGGACCGCTTCTTCAAGACC 60
472 GACATCCGCCAGGGCCCAAGAGAGCCCTTCGGGACTGAGTGGACCGCTTCTTCAAGACC 531
RESULT 9
AAD09488 ID AAD09488 standard; DNA; 1914 BP.
XX AC AAD09488;
XX DT 10-SEP-2001 (first entry)
XX DB Human immunodeficiency virus TA (HIV TA) gene.
XX KW Human immunodeficiency virus TA; HIV TA; immunogen; anti-HIV; vaccine;
XX KW gene therapy; fusion protein; modified vaccinia virus Ankara vector; MVA;
XX KW cytotoxic T-lymphocyte; CTL; epitope; ds.
XX OS Human immunodeficiency virus.
XX FH Key Location/Qualifiers
XX FT CDS 13..1914
XX FT /*tag= a
XX FT /product= "HIV TA immunogen"
XX PN WO200147955-A2.
XX PD 05-JUL-2001.
XX PF 22-DEC-2000; 2000WO-GB004984.
XX PR 23-DEC-1999; 99GB-00030294.
XX PR 14-OCT-2000; 2000GB-00025234.
```

```
XX PA (MEDI-) MEDICAL RES COUNCIL.
XX PA (ITAI-) INT AIDS VACCINE INITIATIVE.
XX PA (UYNA-) UNIV NAIROBI.
XX PI Hanke T, Mcmichael AJ;
XX DR WPI; 2001-418221/44.
XX DR P-PSDB; AAE04826.
XX PT Novel immunogen for stimulating anti-HIV immune response, has a portion
XX PT of gag protein of HIV from HIV clade, parts of p17, p24 and synthetic
XX PT polypeptide comprising human cytotoxic T-lymphocyte epitopes of HIV
XX PT protein.
XX PS Example 2; Fig 6B; 65pp; English.
XX CC The invention relates to human immunodeficiency virus immunogens and
XX CC their corresponding DNA molecules. An immunogen comprises a portion of
XX CC gag protein of HIV from an HIV clade, parts of p17 and p24, modified to
XX CC prevent N-terminal myristoylation; and a synthetic polypeptide comprising
XX CC human cytotoxic T-lymphocyte (CTL) epitopes of HIV protein. This
XX CC immunogen is designed to elicit an HIV-specific immune response in
XX CC humans. The immunogen is useful in the preparation of a medicament such
XX CC as vaccine to prevent or treat HIV infection in a human subject. The
XX CC invention also relates to method of stimulating anti-HIV immune response
XX CC in a human subject which comprises administering one or more times an
XX CC amount of nucleic acid molecule sufficient to prime an immune response to
XX CC the immunogen, or else may be packaged within a delivery means, such as a
XX CC modified vaccinia virus Ankara (MVA) to boost the immune response to
XX CC common portion of the immunogens. The present DNA sequence encodes human
XX CC immunodeficiency virus TA immunogen (HIV TA) fusion protein construct.
XX CC HIV TA construct is based on HIV A immunogen. The HIV A immunogen
XX CC consists of about 73% of gag protein fused to a string of 25 partially
XX CC overlapping human CTL epitopes. The gag domain of HIV A contains p24 and
XX CC p17 in an order reversed to the viral gag p17-p24-p15 polypeptide. HIV TA
XX CC shares the same design rationale with HIV A construct, but additionally
XX CC includes the HIV-1 clade A tat sequence positioned between the gag
XX CC protein and synthetic human CTL epitopes
XX SQ Sequence 1914 BP; 383 A; 728 C; 534 G; 269 T; 0 U; 0 Other;
OY Query Match 100.0%; Score 60; DB 4; Length 1914;
Best Local Similarity 100.0%; Pred. No. 5.8e-09;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1 GACATCCGCCAGGGCCCAAGAGAGCCCTTCGGGACTGAGTGGACCGCTTCTTCAAGACC 60
466 GACATCCGCCAGGGCCCAAGAGAGCCCTTCGGGACTGAGTGGACCGCTTCTTCAAGACC 525
RESULT 10
AAD09489 ID AAD09489 standard; DNA; 2493 BP.
XX AC AAD09489;
XX DT 10-SEP-2001 (first entry)
XX DB Human immunodeficiency virus Aet (HIV Aet) gene.
XX KW Human immunodeficiency virus Aet; HIV Aet; immunogen; anti-HIV; vaccine;
XX KW gene therapy; fusion protein; modified vaccinia virus Ankara vector; MVA;
XX KW cytotoxic T-lymphocyte; CTL; epitope; ds.
XX OS Human immunodeficiency virus.
XX FH Key Location/Qualifiers
XX FT CDS 2179..2493
XX FT /*tag= a
XX FT /product= "HIV Aet immunogen"
XX PN WO200147955-A2.
```

```
XX 05-JUL-2001.
PD 22-DEC-2000; 2000MO-GB004984.
XX 23-DEC-1999; 99GB-00030294.
XX 14-OCT-2000; 2000GB-00025234.
XX (MEDI-) MEDICAL RES COUNCIL.
XX (ITAL-) INT AIDS VACCINE INITIATIVE.
XX (UYNA-) UNIV NAIROBI.
XX Hanke T, Mcmichael AJ;
XX WPI; 2001-418221/44.
XX P-PSDB; AAE04827.
XX Novel immunogen for stimulating anti-HIV immune response, has a portion
XX of gag protein of HIV from HIV clade, parts of p17, p24 and synthetic
XX polypeptide comprising human cytotoxic T-lymphocyte epitopes of HIV
XX protein.
XX Example 2; Fig 7B; 65bp; English.
XX The invention relates to human immunodeficiency virus immunogens and
XX their corresponding DNA molecules. An immunogen comprises a portion of
XX gag protein of HIV from an HIV clade, parts of p17 and p24, modified to
XX prevent N-terminal myristoylation; and a synthetic polypeptide comprising
XX human cytotoxic T-lymphocyte (CTL) epitopes of HIV protein. This
XX immunogen is designed to elicit an HIV-specific immune response in
XX humans. The immunogen is useful in the preparation of a medicament such
XX as vaccine to prevent or treat HIV infection in a human subject. The
XX invention also relates to method of stimulating anti-HIV immune response
XX in a human subject which comprises administering one or more times an
XX amount of nucleic acid molecule sufficient to prime an immune response to
XX the immunogen, or else may be packaged within a delivery means, such as a
XX modified vaccinia virus Ankara (MVA) to boost the immune response to
XX common portion of the immunogens. The present DNA sequence encodes human
XX immunodeficiency virus Aet immunogen (HIV Aet) fusion protein construct.
XX HIV Aet construct is based on HIV A immunogen. The HIV A immunogen
XX consists of about 73% of gag protein fused to a string of 25 partially
XX overlapping human CTL epitopes. The gag domain of HIV A contains p24 and
XX p17 in an order reversed to the viral gag p17-p24-p15 polypeptide. HIV
XX Aet shares the same design rationale with HIV A construct, but
XX additionally includes the HIV-1 clade A tat sequence which is present on
XX the same construct, but expressed as a separate polypeptide, by virtue of
XX the inclusion of an internal ribosome entry site
XX
XX Sequence 2493 BP; 516 A; 887 C; 685 G; 405 T; 0 U; 0 Other;
SQ
XX Query Match 100.0%; Score 60; DB 4; Length 2493;
XX Best Local Similarity 100.0%; Pred. No. 5.9e-09;
XX Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GACATCCGCCAGGCCCCCAAGAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 60
DB 466 GACATCCGCCAGGCCCCCAAGAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 525
RESULT 11
AAD09490
ID AAD09490 standard; DNA; 4350 BP.
AC AAD09490;
XX 10-SEP-2001 (first entry)
XX Human immunodeficiency virus PPA immunogen.
XX Human immunodeficiency virus; PPA; immunogen; anti-HIV; vaccine;
XX gene therapy; fusion protein; modified vaccinia virus Ankara vector; MVA;
XX cytotoxic T-lymphocyte; CTL; epitope; ds.
```

```
OS Human immunodeficiency virus.
XX Key Location/Qualifiers
XX CDS 13..4350
XX FT /*tag= a
XX FT /product= "PPA immunogen"
XX WO200147955-A2.
XX 05-JUL-2001.
XX 22-DEC-2000; 2000MO-GB004984.
XX 23-DEC-1999; 99GB-00030294.
XX 14-OCT-2000; 2000GB-00025234.
XX (MEDI-) MEDICAL RES COUNCIL.
XX (ITAL-) INT AIDS VACCINE INITIATIVE.
XX (UYNA-) UNIV NAIROBI.
XX Hanke T, Mcmichael AJ;
XX WPI; 2001-418221/44.
XX P-PSDB; AAE04828.
XX Novel immunogen for stimulating anti-HIV immune response, has a portion
XX of gag protein of HIV from HIV clade, parts of p17, p24 and synthetic
XX polypeptide comprising human cytotoxic T-lymphocyte epitopes of HIV
XX protein.
XX Claim 40; Fig 8B; 65bp; English.
XX The invention relates to human immunodeficiency virus immunogens and
XX their corresponding DNA molecules. An immunogen comprises a portion of
XX gag protein of HIV from an HIV clade, parts of p17 and p24, modified to
XX prevent N-terminal myristoylation; and a synthetic polypeptide comprising
XX human cytotoxic T-lymphocyte (CTL) epitopes of HIV protein. This
XX immunogen is designed to elicit an HIV-specific immune response in
XX humans. The immunogen is useful in the preparation of a medicament such
XX as vaccine to prevent or treat HIV infection in a human subject. The
XX invention also relates to method of stimulating anti-HIV immune response
XX in a human subject which comprises administering one or more times an
XX amount of nucleic acid molecule sufficient to prime an immune response to
XX the immunogen, or else may be packaged within a delivery means, such as a
XX modified vaccinia virus Ankara (MVA) to boost the immune response to
XX common portion of the immunogens. The present DNA sequence encodes human
XX immunodeficiency virus PPA immunogen fusion protein construct. PPA
XX construct is based on HIV A immunogen. The HIV A immunogen consists of
XX about 73% of gag protein fused to a string of 25 partially overlapping
XX human CTL epitopes. The gag domain of HIV A contains p24 and p17 in an
XX order reversed to the viral gag p17-p24-p15 polypeptide. PPA shares the
XX same design rationale with HIV A construct, but additionally includes the
XX HIV-1 clade A tat sequence, C-pol protein, N-pol protein, rev and nef
XX proteins
XX
XX Sequence 4350 BP; 888 A; 1630 C; 1259 G; 573 T; 0 U; 0 Other;
SQ
XX Query Match 100.0%; Score 60; DB 4; Length 4350;
XX Best Local Similarity 100.0%; Pred. No. 6.1e-09;
XX Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GACATCCGCCAGGCCCCCAAGAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 60
DB 466 GACATCCGCCAGGCCCCCAAGAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 525
RESULT 12
AAA70426
ID AAA70426 standard; DNA; 60 BP.
XX AAA70426;
XX 15-SEP-2003 (revised)
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 28, 2004, 09:05:41 ; Search time 246.567 Seconds

(without alignment)  
10547.179 Million cell updates/sec

Title: US-09-475-704A-2

Perfect score: 60  
Sequence: 1 gacatccgcgcagggccca.....tgacccgtcttcaagacc 60

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenBml: \*  
1: gb\_ba: \*  
2: gb\_htg: \*  
3: gb\_in: \*  
4: gb\_om: \*  
5: gb\_ov: \*  
6: gb\_pat: \*  
7: gb\_ph: \*  
8: gb\_pl: \*  
9: gb\_pr: \*  
10: gb\_ro: \*  
11: gb\_sts: \*  
12: gb\_sy: \*  
13: gb\_un: \*  
14: gb\_vi: \*  
15: em\_ba: \*  
16: em\_fun: \*  
17: em\_hum: \*  
18: em\_in: \*  
19: em\_mu: \*  
20: em\_om: \*  
21: em\_or: \*  
22: em\_ov: \*  
23: em\_pat: \*  
24: em\_ph: \*  
25: em\_pl: \*  
26: em\_ro: \*  
27: em\_sts: \*  
28: em\_un: \*  
29: em\_vi: \*  
30: em\_htg\_hum: \*  
31: em\_htg\_inv: \*  
32: em\_htg\_other: \*  
33: em\_htg\_mus: \*  
34: em\_htg\_pln: \*  
35: em\_htg\_rod: \*  
36: em\_htg\_man: \*  
37: em\_htg\_vrt: \*  
38: em\_sy: \*  
39: em\_htgo\_hum: \*  
40: em\_htgo\_mus: \*  
41: em\_htgo\_other: \*

Score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	60	100.0	60	AX455886	AX455886 Sequence
2	60	100.0	1509	AX455888	AX455888 Sequence
3	60	100.0	1509	AX455905	AX455905 Sequence
4	60	100.0	1509	AX468544	AX468544 Sequence
5	60	100.0	1509	AX468548	AX468548 Sequence
6	60	100.0	1608	AX188558	AX188558 Sequence
7	60	100.0	1914	AX188560	AX188560 Sequence
8	60	100.0	2493	AX188562	AX188562 Sequence
9	60	100.0	4350	AX188564	AX188564 Sequence
10	58.4	97.3	60	BD263651	BD263651 Improved
11	58.4	97.3	60	AR373334	AR373334 Sequence
12	58.4	97.3	1268	BD263642	BD263642 Improved
13	58.4	97.3	1268	AR373325	AR373325 Sequence
14	58.4	97.3	1509	AF201927	AF201927 Synthetic
15	58.4	97.3	1515	BD263637	BD263637 Improved
16	58.4	97.3	1515	AR373320	AR373320 Sequence
17	58.4	97.3	1847	AF202464	AF202464 Synthetic
18	58.4	97.3	1847	AF202465	AF202465 Synthetic
19	58.4	97.3	1853	BD263638	BD263638 Improved
20	58.4	97.3	1853	AR373321	AR373321 Sequence
21	58.4	97.3	1865	BD263700	BD263700 Improved
22	58.4	97.3	1865	BD263701	BD263701 Improved
23	58.4	97.3	1865	AR373383	AR373383 Sequence
24	58.4	97.3	1865	AR373384	AR373384 Sequence
25	58.4	97.3	2031	BD263640	BD263640 Improved
26	58.4	97.3	2031	AR373323	AR373323 Sequence
27	58.4	97.3	4319	BD263639	BD263639 Improved
28	58.4	97.3	4319	AR373322	AR373322 Sequence
29	58.4	97.3	4472	BD263697	BD263697 Improved
30	58.4	97.3	4472	AR373380	AR373380 Sequence
31	58.4	97.3	4608	BD263698	BD263698 Improved
32	58.4	97.3	4608	AR373381	AR373381 Sequence
33	58.4	97.3	4689	BD263696	BD263696 Improved
34	58.4	97.3	4689	AR373379	AR373379 Sequence
35	58.4	97.3	4766	BD263695	BD263695 Improved
36	58.4	97.3	4766	AR373378	AR373378 Sequence
37	58.4	97.3	6438	AX427919	AX427919 Sequence
38	58.4	97.3	9167	AX427933	AX427933 Sequence
39	58.4	97.3	9170	AX427928	AX427928 Sequence
40	58.4	97.3	9189	AX427921	AX427921 Sequence
41	58.4	97.3	9194	AX427925	AX427925 Sequence
42	58.4	97.3	9194	AX427926	AX427926 Sequence
43	58.4	97.3	9407	AX427937	AX427937 Sequence
44	58.4	97.3	9782	AX427935	AX427935 Sequence
45	58.4	97.3	9783	AX427934	AX427934 Sequence

ALIGNMENTS

RESULT 1  
LOCUS AX455886 60 bp DNA linear PAT 06-JUL-2002  
DEFINITION Sequence 2 from Patent WO0204493.  
ACCESSION AX455886  
VERSION AX455886.1 GI:21714879  
KEYWORDS  
SOURCE Human immunodeficiency virus  
ORGANISM Human immunodeficiency virus  
REFERENCE 1  
AUTHORS zur Megede, J., Barnett, S.W., Engelbrecht, S. and van Rensburg, R.  
TITLE Polynucleotides encoding antigenic hiv type c polypeptides,  
polypeptides and uses thereof

Pred. No. is the number of results predicted by chance to have a

JOURNAL Patent: WO 0204493-A 2 17-JAN-2002;  
CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)  
FEATURES  
source  
1. 60  
/organism="Human immunodeficiency virus"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:12721"

ORIGIN  
Query Match 100.0%; Score 60; DB 6; Length 60;  
Best Local Similarity 100.0%; Pred. No. 9.7e-07;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACATCCGCGAGGGCCCCAAGAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 60  
1 GACATCCGCGAGGGCCCCAAGAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 60

Db 1 GACATCCGCGAGGGCCCCAAGAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 60

RESULT 2  
AX455888 1509 bp DNA linear PAT 06-JUL-2002  
LOCUS Sequence 4 from Patent WO0204493.  
ACCESSION AX455888  
VERSION AX455888.1 GI:21714881  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1  
AUTHORS zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.  
TITLE Polynucleotides encoding antigenic hiv type c polypeptides,  
JOURNAL Patent: WO 0204493-A 4 17-JAN-2002;  
CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)  
FEATURES  
source  
1. 1509  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="synthetic Gag of HIV strain AF110967"

ORIGIN  
Query Match 100.0%; Score 60; DB 6; Length 1509;  
Best Local Similarity 100.0%; Pred. No. 5.6e-07;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACATCCGCGAGGGCCCCAAGAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 60  
1 GACATCCGCGAGGGCCCCAAGAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 60

Db 841 GACATCCGCGAGGGCCCCAAGAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 900

RESULT 3  
AX455905 1509 bp DNA linear PAT 06-JUL-2002  
LOCUS Sequence 21 from Patent WO0204493.  
ACCESSION AX455905  
VERSION AX455905.1 GI:21714897  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1  
AUTHORS zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.  
TITLE Polynucleotides encoding antigenic hiv type c polypeptides,  
JOURNAL Patent: WO 0204493-A 21 17-JAN-2002;  
CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)  
FEATURES  
source  
1. 1509  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="synthetic Gag coding sequence of HIV strain

AF110967"

ORIGIN  
Query Match 100.0%; Score 60; DB 6; Length 1509;  
Best Local Similarity 100.0%; Pred. No. 5.6e-07;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACATCCGCGAGGGCCCCAAGAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 60  
1 GACATCCGCGAGGGCCCCAAGAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 60

Db 841 GACATCCGCGAGGGCCCCAAGAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 900

RESULT 4  
AX468544 1509 bp DNA linear PAT 16-JUL-2002  
LOCUS Sequence 64 from Patent WO0226209.  
ACCESSION AX468544  
VERSION AX468544.1 GI:21901374  
KEYWORDS  
SOURCE Human immunodeficiency virus 1 (HIV-1)  
ORGANISM Human immunodeficiency virus 1  
REFERENCE 1  
AUTHORS O'Hagan,D., Otten,G., Donnelly,J.J., Polo,J.M., Barnett,S.,  
TITLE Singh,M., Ulmer,J. and Dubensky,T.W.  
JOURNAL Metoparticles for delivery of the heterologous nucleic acids  
Patent: WO 0226209-A 64 04-APR-2002;  
CHIRON CORPORATION (US)  
FEATURES  
source  
1. 1509  
/organism="Human immunodeficiency virus 1"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:11676"

ORIGIN  
Query Match 100.0%; Score 60; DB 6; Length 1509;  
Best Local Similarity 100.0%; Pred. No. 5.6e-07;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACATCCGCGAGGGCCCCAAGAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 60  
1 GACATCCGCGAGGGCCCCAAGAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 60

Db 841 GACATCCGCGAGGGCCCCAAGAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 900

RESULT 5  
AX468548 1509 bp DNA linear PAT 16-JUL-2002  
LOCUS Sequence 68 from Patent WO0226209.  
ACCESSION AX468548  
VERSION AX468548.1 GI:21901378  
KEYWORDS  
SOURCE Human immunodeficiency virus 1 (HIV-1)  
ORGANISM Human immunodeficiency virus 1  
REFERENCE 1  
AUTHORS O'Hagan,D., Otten,G., Donnelly,J.J., Polo,J.M., Barnett,S.,  
TITLE Singh,M., Ulmer,J. and Dubensky,T.W.  
JOURNAL Microparticles for delivery of the heterologous nucleic acids  
Patent: WO 0226209-A 68 04-APR-2002;  
CHIRON CORPORATION (US)  
FEATURES  
source  
1. 1509  
/organism="Human immunodeficiency virus 1"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:11676"

ORIGIN  
Query Match 100.0%; Score 60; DB 6; Length 1509;  
Best Local Similarity 100.0%; Pred. No. 5.6e-07;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GACATCCGCGCAGGGCCCGCAGAGAGCCCTTCCGCGACTAGCTGAGACCGCTTCTCAAGACC 60  
DB 841 GACATCCGCGCAGGGCCCGCAGAGAGCCCTTCCGCGACTAGCTGAGACCGCTTCTCAAGACC 900

RESULT 6  
AXI88558  
LOCUS AXI88558 1608 bp DNA linear PAT 08-AUG-2001  
DEFINITION Sequence 27 from Patent WO0147955.  
ACCESSION AXI88558  
VERSION AXI88558.1 GI:15142199  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
artificial sequences.

REFERENCE  
1 Hanke,T.M. and Mcmichael,A.J.  
AUTHORS Improvements in or relating to immune responses to hiv  
TITLE Patent: WO 0147955-A 27 05-JUL-2001;  
JOURNAL MEDICAL RESEARCH COUNCIL (GB) ; International Aids Vaccine  
Initiative (US) ; University of Nairobi (KE)  
LOCATION/Qualifiers

FEATURES  
SOURCE 1.1608  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="Chimeric polynucleotide"

ORIGIN

Query Match 100.0%; Score 60; DB 6; Length 1608;  
Best Local Similarity 100.0%; Pred. No. 5.5e-07;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GACATCCGCGCAGGGCCCGCAGAGAGCCCTTCCGCGACTAGCTGAGACCGCTTCTCAAGACC 60  
DB 472 GACATCCGCGCAGGGCCCGCAGAGAGCCCTTCCGCGACTAGCTGAGACCGCTTCTCAAGACC 531

RESULT 7  
AXI88560  
LOCUS AXI88560 1914 bp DNA linear PAT 08-AUG-2001  
DEFINITION Sequence 29 from Patent WO0147955.  
ACCESSION AXI88560  
VERSION AXI88560.1 GI:15142200  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
artificial sequences.

REFERENCE  
1 Hanke,T.M. and Mcmichael,A.J.  
AUTHORS Improvements in or relating to immune responses to hiv  
TITLE Patent: WO 0147955-A 29 05-JUL-2001;  
JOURNAL MEDICAL RESEARCH COUNCIL (GB) ; International Aids Vaccine  
Initiative (US) ; University of Nairobi (KE)  
LOCATION/Qualifiers

FEATURES  
SOURCE 1.1914  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="Chimeric polynucleotide"

ORIGIN

Query Match 100.0%; Score 60; DB 6; Length 1914;  
Best Local Similarity 100.0%; Pred. No. 5.4e-07;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GACATCCGCGCAGGGCCCGCAGAGAGCCCTTCCGCGACTAGCTGAGACCGCTTCTCAAGACC 60  
DB 466 GACATCCGCGCAGGGCCCGCAGAGAGCCCTTCCGCGACTAGCTGAGACCGCTTCTCAAGACC 525

RESULT 8  
AXI88562

LOCUS AXI88562 2493 bp DNA linear PAT 08-AUG-2001  
DEFINITION Sequence 31 from Patent WO0147955.  
ACCESSION AXI88562  
VERSION AXI88562.1 GI:15142201  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
artificial sequences.

REFERENCE  
1 Hanke,T.M. and Mcmichael,A.J.  
AUTHORS Improvements in or relating to immune responses to hiv  
TITLE Patent: WO 0147955-A 31 05-JUL-2001;  
JOURNAL MEDICAL RESEARCH COUNCIL (GB) ; International Aids Vaccine  
Initiative (US) ; University of Nairobi (KE)  
LOCATION/Qualifiers

FEATURES  
SOURCE 1.2493  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="Chimeric polynucleotide"

ORIGIN

Query Match 100.0%; Score 60; DB 6; Length 2493;  
Best Local Similarity 100.0%; Pred. No. 5.1e-07;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GACATCCGCGCAGGGCCCGCAGAGAGCCCTTCCGCGACTAGCTGAGACCGCTTCTCAAGACC 60  
DB 466 GACATCCGCGCAGGGCCCGCAGAGAGCCCTTCCGCGACTAGCTGAGACCGCTTCTCAAGACC 525

RESULT 9  
AXI88564  
LOCUS AXI88564 4350 bp DNA linear PAT 08-AUG-2001  
DEFINITION Sequence 33 from Patent WO0147955.  
ACCESSION AXI88564  
VERSION AXI88564.1 GI:15142202  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
artificial sequences.

REFERENCE  
1 Hanke,T.M. and Mcmichael,A.J.  
AUTHORS Improvements in or relating to immune responses to hiv  
TITLE Patent: WO 0147955-A 33 05-JUL-2001;  
JOURNAL MEDICAL RESEARCH COUNCIL (GB) ; International Aids Vaccine  
Initiative (US) ; University of Nairobi (KE)  
LOCATION/Qualifiers

FEATURES  
SOURCE 1.4350  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="Chimeric polynucleotide"

ORIGIN

Query Match 100.0%; Score 60; DB 6; Length 4350;  
Best Local Similarity 100.0%; Pred. No. 4.7e-07;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GACATCCGCGCAGGGCCCGCAGAGAGCCCTTCCGCGACTAGCTGAGACCGCTTCTCAAGACC 60  
DB 466 GACATCCGCGCAGGGCCCGCAGAGAGCCCTTCCGCGACTAGCTGAGACCGCTTCTCAAGACC 525

RESULT 10  
BD263651  
LOCUS BD263651 60 bp DNA linear PAT 17-JUL-2003  
DEFINITION Improved expression of HIV polypeptides and production of  
virus-like particles.  
ACCESSION BD263651  
VERSION BD263651.1 GI:33073419  
KEYWORDS JP 2002533124-A/18.  
SOURCE synthetic construct



ORGANISM synthetic construct  
artificial sequences.  
REFERENCE 1 (bases 1 to 60)  
AUTHORS Barnett,S., Megede,J.Z., Srivastava,I., Lian,Y., Hartog,K., Liu,H., Greer,C., Selby,M. and Walker,C.  
TITLE Improved expression of HIV polypeptides and production of virus-like particles  
JOURNAL Patent: JP 2002533124-A 18 08-OCT-2002;  
CHIRON CORP  
COMMENT OS Artificial Sequence  
PN JP 2002533124-A/18  
PD 08-OCT-2002  
PF 30-DEC-1999 JP 2000591193  
PR 31-DEC-1998 US 60/114495,01-DEC-1999 US 60/168471 PI  
SUSAN BARNETT,JAN ZUR MEGEDE,INDRESH SRIVASTAVA,YING LIAN, PI  
KARIN HARTOG,  
PI HONG LIU,CATHERINE GREER,MARK SELBY,CHRISTOPHER WALKER PC  
C12N15/09,A61K31/711,A61K38/00,A61K48/00,A61P31/18,A61P37/02, PC  
C12N5/10,  
PC C12N7/00,C12P21/02,C12N15/00,C12N5/00,A61K37/02 CC  
Description of Artificial Sequence: synthetic p55 Gag CC Major  
Homology Region  
FH Key Location/Qualifiers  
FT source 1..60  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"

ORIGIN  
Query Match 97.3%; Score 58.4; DB 6; Length 60;  
Best Local Similarity 98.3%; Pred. No. 2.7e-06;  
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GACATCCGCGCAGGGCCCCCAAGAGAGCCCTTCCGCGACTACGTGACCGCTTCTCAAGACC 60  
Db 1 GACATCCGCGCAGGGCCCCCAAGAGAGCCCTTCCGCGACTACGTGACCGCTTCTCAAGACC 60

RESULT 11  
AR373334 60 bp DNA linear PAT 18-DEC-2003  
LOCUS AR373334  
DEFINITION Sequence 20 from patent US 6602705.  
ACCESSION AR373334  
VERSION AR373334.1 GI:40075437  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 60)  
AUTHORS Barnett,S.W., Megede,J., Greer,C. and Selby,M.  
TITLE Expression of HIV polypeptides and production of virus-like particles  
JOURNAL Patent: US 6602705-A 20 05-AUG-2003;  
FEATURES  
LOCATION/Qualifiers  
SOURCE 1..60  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 97.3%; Score 58.4; DB 6; Length 60;  
Best Local Similarity 98.3%; Pred. No. 2.7e-06;  
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GACATCCGCGCAGGGCCCCCAAGAGAGCCCTTCCGCGACTACGTGACCGCTTCTCAAGACC 60  
Db 1 GACATCCGCGCAGGGCCCCCAAGAGAGCCCTTCCGCGACTACGTGACCGCTTCTCAAGACC 60

RESULT 12  
BD263642

LOCUS BD263642 1268 bp DNA linear PAT 17-JUL-2003  
DEFINITION Improved expression of HIV polypeptides and production of virus-like particles.  
ACCESSION BD263642  
VERSION BD263642.1 GI:33073410  
KEYWORDS JP 2002533124-A/9.  
SOURCE JP 2002533124-A/9.  
ORGANISM synthetic construct  
artificial sequences.  
REFERENCE 1 (bases 1 to 1268)  
AUTHORS Barnett,S., Megede,J.Z., Srivastava,I., Lian,Y., Hartog,K., Liu,H., Greer,C., Selby,M. and Walker,C.  
TITLE Improved expression of HIV polypeptides and production of virus-like particles  
JOURNAL Patent: JP 2002533124-A 9 08-OCT-2002;  
CHIRON CORP  
COMMENT OS Artificial Sequence  
PN JP 2002533124-A/9  
PD 08-OCT-2002  
PF 30-DEC-1999 JP 2000591193  
PR 31-DEC-1998 US 60/114495,01-DEC-1999 US 60/168471 PI  
SUSAN BARNETT,JAN ZUR MEGEDE,INDRESH SRIVASTAVA,YING LIAN, PI  
KARIN HARTOG,  
PI HONG LIU,CATHERINE GREER,MARK SELBY,CHRISTOPHER WALKER PC  
C12N15/09,A61K31/711,A61K38/00,A61K48/00,A61P31/18,A61P37/02, PC  
C12N5/10,  
PC C12N7/00,C12P21/02,C12N15/00,C12N5/00,A61K37/02 CC  
Description of Artificial Sequence: synthetic Gag common CC  
region  
FH Key Location/Qualifiers  
FT source 1..1268  
/organism="Artificial Sequence".  
/location/Qualifiers  
1..1268  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"

ORIGIN  
Query Match 97.3%; Score 58.4; DB 6; Length 1268;  
Best Local Similarity 98.3%; Pred. No. 1.6e-06;  
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GACATCCGCGCAGGGCCCCCAAGAGAGCCCTTCCGCGACTACGTGACCGCTTCTCAAGACC 60  
Db 862 GACATCCGCGCAGGGCCCCCAAGAGAGCCCTTCCGCGACTACGTGACCGCTTCTCAAGACC 921

RESULT 13  
AR373325 1268 bp DNA linear PAT 18-DEC-2003  
LOCUS AR373325  
DEFINITION Sequence 9 from patent US 6602705.  
ACCESSION AR373325  
VERSION AR373325.1 GI:40075428  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1268)  
AUTHORS Barnett,S.W., Megede,J., Greer,C. and Selby,M.  
TITLE Expression of HIV polypeptides and production of virus-like particles  
JOURNAL Patent: US 6602705-A 9 05-AUG-2003;  
FEATURES  
LOCATION/Qualifiers  
SOURCE 1..1268  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 97.3%; Score 58.4; DB 6; Length 1268;  
Best Local Similarity 98.3%; Pred. No. 1.6e-06;  
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GACATCCGCCAGGGCCCCCAAGAGAGCCCTTCCGCGACTAGTGGACCGGCTTCTCAAGACC 60  
 |||  
 DB 862 GACATCCGCCAGGGCCCCCAAGAGAGCCCTTCCGCGACTAGTGGACCGGCTTCTCAAGACC 921

## RESULT 14

AF201927 1509 bp DNA linear SYN 16-MAR-2000  
 LOCUS AF201927  
 DEFINITION Synthetic construct gag protein gene, complete cds.  
 ACCESSION AF201927  
 VERSION AF201927.1 GI:7248702

KEYWORDS  
 SOURCE synthetic construct  
 ORGANISM synthetic construct

REFERENCE 1 (bases 1 to 1509)  
 AUTHORS zur Megede,J., Chen,M.C., Doe,B., Schaefer,M., Greer,C.E.,  
 Selby,M., Otten,G.R. and Barnett,S.W.  
 TITLE Increased expression and immunogenicity of sequence-modified human  
 JOURNAL immunodeficiency virus type 1 gag gene  
 MEDLINE J. Virol. 74 (6), 2628-2635 (2000)  
 PUBMED 20148954  
 10684277

REFERENCE 2 (bases 1 to 1509)  
 AUTHORS zur Megede,J. and Barnett,S.W.  
 TITLE Direct Submission  
 JOURNAL Submitted (04-NOV-1999) Vaccines, Chiron Corporation, 4560 Horton,  
 Emeryville, CA 94608, USA

FEATURES  
 source Location/Qualifiers

1..1509  
 /organism="synthetic construct"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:32630"  
 1..1509  
 /note="HIV-1SF2 p55gag"  
 /codon\_start=1  
 /transl\_table=1  
 /product="gag protein"  
 /protein\_id="AAF43628.1"  
 /db\_xref="GI:7248703"

/translation="MGARASVSGGELDKWEKIRLRPGKKKYLKHIVASRELERF  
 AVNGLBTSBGRQILGQLPSLQTSBELSLNTVATLYCVHORIDVKDTKALE  
 KIEEONKSKKKAQAAAGTGNSSOVSNYPVQNQGOMVQAISPTLNAMVKY  
 VBEKAFSPBVLPMFSALSBGATPDQDLNTMLNTVGGHQAAMQMLKETINEBAENDRVH  
 PVHAGPIAPGQMRBPRSDIAGTSTLQBOIGMTNPNPIPVGBIYKRWIILGLNKIV  
 RMYSPSTILDIRQPKPEPRDYVDRFYKTLRAEQASQDVANMTETLLVQANPDCKT  
 ILKALGPAATLEMTACOGVGGHKAARVLAEMSOVTNPATIMORGNFNRQRTV  
 KCFNCGKEGHTARNCRAPRKCKGRCGRBGHQMDCTERQANFLGKIWPSTYKGRPGNF  
 LQSRPEPTAPPEBSFRFGBEKTPPSQKQBPIDKELYPILTSLRSLFGNDPSSQ"

## ORIGIN

Query Match 97.3%; Score 58.4; DB 12; Length 1509;  
 Best Local Similarity 98.3%; Pred. No. 1.6e-06;  
 Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GACATCCGCCAGGGCCCCCAAGAGAGCCCTTCCGCGACTAGTGGACCGGCTTCTCAAGACC 60  
 |||  
 DB 862 GACATCCGCCAGGGCCCCCAAGAGAGCCCTTCCGCGACTAGTGGACCGGCTTCTCAAGACC 915

## RESULT 15

BD263637 1515 bp DNA linear PAT 17-JUL-2003  
 LOCUS BD263637  
 DEFINITION Improved expression of HIV polypeptides and production of  
 virus-like particles.

ACCESSION BD263637  
 VERSION BD263637.1 GI:33073405  
 KEYWORDS JP 2002533124-A/4.  
 SOURCE synthetic construct  
 ORGANISM synthetic construct

REFERENCE 1 (bases 1 to 1515)  
 AUTHORS Barnett,S., Megede,J.Z., Srivastava,I., Lian,Y., Hartog,K., Liu,H.,

TITLE Greer,C., Selby,M. and Walker,C.  
 JOURNAL Improved expression of HIV polypeptides and production of  
 virus-like particles  
 COMMENT Patent: JP 2002533124-A 4 08-OCT-2002;  
 CHIRON CORP

OS Artificial Sequence  
 PN JP 2002533124-A/4  
 PD 08-OCT-2002  
 PF 30-DEC-1999 JP 2000591193  
 PR 31-DEC-1998 US 60/114495, 01-DEC-1999 US 60/168471 PI  
 SUSAN BARNETT, JAN ZUR MEGEDE, INDRESH SRIVASTAVA, YING LIAN, PI  
 KARIN HARTOG,

PI HONG LIU, CATHERINE GREER, MARK SELBY, CHRISTOPHER WALKER PC  
 C12N15/09, A61K31/711, A61K38/00, A61K48/00, A61P31/18, A61P37/02, PC  
 C12N5/10,  
 PC C12N7/00, C12P21/02, C12N15/00, C12N5/00, A61K37/02 CC  
 Description of Artificial Sequence: synthetic HIV-Gag FH Key

FEATURES  
 source Location/Qualifiers

FT source 1..1515  
 /organism="Artificial Sequence".  
 1..1515  
 /organism="synthetic construct"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:32630"

## ORIGIN

Query Match 97.3%; Score 58.4; DB 6; Length 1515;  
 Best Local Similarity 98.3%; Pred. No. 1.6e-06;  
 Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GACATCCGCCAGGGCCCCCAAGAGAGCCCTTCCGCGACTAGTGGACCGGCTTCTCAAGACC 60  
 |||  
 DB 862 GACATCCGCCAGGGCCCCCAAGAGAGCCCTTCCGCGACTAGTGGACCGGCTTCTCAAGACC 921

Search completed: May 28, 2004, 13:23:16  
 Job time : 247.567 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 28, 2004, 11:33:50 ; Search time 30.7692 Seconds  
(without alignments)  
8870.358 Million cell updates/sec

Title: US-09-475-704A-1  
Perfect score: 60  
Sequence: 1 gacatcaagcaggcccccacaa.....tggaccgccttctcaagacc 60

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2960401 seqs, 2274450654 residues

Total number of hits satisfying chosen parameters: 5920802

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_MA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 18: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 19: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	13	US-09-899-575-1
2	60	100.0	60	13	US-09-899-575-53
3	60	100.0	1479	10	US-09-967-464-63
4	60	100.0	1479	10	US-09-967-464-67
5	60	100.0	1479	13	US-09-899-575-3
6	60	100.0	1479	13	US-09-899-575-20
7	60	100.0	1491	13	US-09-899-575-99
8	60	100.0	1494	13	US-09-899-575-51
9	60	100.0	2742	15	US-10-190-435-20
10	60	100.0	2742	15	US-10-190-435-57
11	60	100.0	2742	15	US-10-190-305A-15
12	60	100.0	2742	15	US-10-190-305A-81
13	60	100.0	3162	15	US-10-190-435-18
14	60	100.0	3462	15	US-10-190-435-16

15	60	100.0	3531	15	US-10-190-435-13	Sequence 13, Appl
16	60	100.0	3537	15	US-10-190-435-14	Sequence 14, Appl
17	60	100.0	3537	15	US-10-190-435-15	Sequence 15, Appl
18	60	100.0	3930	15	US-10-190-435-9	Sequence 9, Appl1
19	60	100.0	3930	15	US-10-190-435-10	Sequence 10, Appl
20	60	100.0	3930	15	US-10-190-435-11	Sequence 11, Appl
21	60	100.0	4419	15	US-10-190-435-19	Sequence 19, Appl
22	60	100.0	4419	15	US-10-190-305A-14	Sequence 14, Appl
23	60	100.0	4483	15	US-10-190-435-35	Sequence 35, Appl
24	60	100.0	4606	15	US-10-190-435-34	Sequence 34, Appl
25	60	100.0	4615	15	US-10-190-435-36	Sequence 36, Appl
26	60	100.0	4702	15	US-10-190-435-38	Sequence 38, Appl
27	60	100.0	4713	15	US-10-190-435-59	Sequence 59, Appl
28	60	100.0	4713	15	US-10-190-305A-83	Sequence 83, Appl
29	60	100.0	4716	15	US-10-190-435-17	Sequence 17, Appl
30	60	100.0	5145	15	US-10-190-435-12	Sequence 12, Appl
31	60	100.0	5145	15	US-10-190-435-13	Sequence 13, Appl
32	60	100.0	5145	15	US-10-190-305A-12	Sequence 12, Appl
33	60	100.0	5184	15	US-10-190-435-58	Sequence 58, Appl
34	60	100.0	5184	15	US-10-190-305A-82	Sequence 82, Appl
35	60	100.0	9166	13	US-10-359-120-168	Sequence 168, App
36	60	100.0	9169	13	US-10-359-120-169	Sequence 169, App
37	60	100.0	9785	13	US-10-359-120-176	Sequence 176, App
38	60	100.0	9788	13	US-10-359-120-174	Sequence 174, App
39	58.4	97.3	1092	15	US-10-138-098-12	Sequence 12, Appl
40	58.4	97.3	1092	15	US-10-138-098-13	Sequence 13, Appl
41	58.4	97.3	3020	15	US-10-138-098-21	Sequence 21, Appl
42	58.4	97.3	3021	15	US-10-138-098-22	Sequence 22, Appl
43	58.4	97.3	3021	15	US-10-138-098-23	Sequence 23, Appl
44	58.4	97.3	3021	15	US-10-138-098-24	Sequence 24, Appl
45	55.2	92.0	60	13	US-09-899-575-2	Sequence 2, Appl1

## ALIGNMENTS

RESULT 1  
US-09-899-575-1  
; Sequence 1, Application US/09899575  
; Publication No. US20030223961A1  
; GENERAL INFORMATION:  
; APPLICANT: Zur Megede, Jan  
; APPLICANT: Barnett, Susan W.  
; APPLICANT: Egnelbrecht, Susan  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C  
; FILE REFERENCE: PP01631.102  
; CURRENT APPLICATION NUMBER: US/09/899,575  
; CURRENT FILING DATE: 2001-07-05  
; PRIOR APPLICATION NUMBER: 09/475,704  
; PRIOR FILING DATE: 1999-12-30  
; NUMBER OF SEQ ID NOS: 135  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 60  
; TYPE: DNA  
; ORGANISM: Human immunodeficiency virus  
; US-09-899-575-1

Query Match 100.0%; Score 60; DB 13; Length 60;  
Best Local Similarity 100.0%; Pred. No. 2.5e-12;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATCAAGCAGGGCCCAAGAGCCCTTCGCGACTACGTGACCGCTTCTCAAGACC 60  
Db 1 GACATCAAGCAGGGCCCAAGAGCCCTTCGCGACTACGTGACCGCTTCTCAAGACC 60

RESULT 2  
US-09-899-575-53  
; Sequence 53, Application US/09899575  
; Publication No. US20030223961A1

```

; GENERAL INFORMATION:
; APPLICANT: Zur Megede, Jan
; APPLICANT: Barnett, Susan W.
; APPLICANT: Egnelbrecht, Susan
; APPLICANT: van Rensburg, Estrelita Janse
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP01631.102
; CURRENT APPLICATION NUMBER: US/09/899,575
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 09/475,704
; PRIOR FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 53
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: HIV Type C Gag
; OTHER INFORMATION: Major Homology Region Optimized
US-09-899-575-53

```

```

Query Match          100.0%; Score 60; DB 13; Length 60;
Best Local Similarity 100.0%; Pred. No. 2.6e-12;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 GACATCAAGCAGGCGCCCAAGAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 60
        |||
        1 GACATCAAGCAGGCGCCCAAGAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 60
DB

```

#### RESULT 3

```

US-09-967-464-63
; Sequence 63, Application US/09967464
; Publication No. US20030138453A1
; GENERAL INFORMATION:

```

```

; APPLICANT: O'Hagan, Derek
; APPLICANT: Otten, Gillis
; APPLICANT: Donnelly, John J.
; APPLICANT: Polo, John M.
; APPLICANT: Barnett, Susan
; APPLICANT: Singh, Manohan
; APPLICANT: Ulmer, Jeffrey
; APPLICANT: Dubensky, Jr., Thomas W.
; TITLE OF INVENTION: MICROPARTICLES FOR DELIVERY OF HETEROLOGOUS NUCLEIC ACIDS
; FILE REFERENCE: PP16269.004
; CURRENT APPLICATION NUMBER: US/09/967,464
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 60/236,105
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 60/315,905
; PRIOR FILING DATE: 2001-08-30
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 63
; LENGTH: 1479
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
US-09-967-464-63

```

```

Query Match          100.0%; Score 60; DB 10; Length 1479;
Best Local Similarity 100.0%; Pred. No. 2.6e-12;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 GACATCAAGCAGGCGCCCAAGAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 60
        |||
        844 GACATCAAGCAGGCGCCCAAGAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 903
DB

```

RESULT 4  
US-09-967-464-67  
; Sequence 67, Application US/09967464

```

; Publication No. US20030138453A1
; GENERAL INFORMATION:
; APPLICANT: O'Hagan, Derek
; APPLICANT: Otten, Gillis
; APPLICANT: Donnelly, John J.
; APPLICANT: Polo, John M.
; APPLICANT: Barnett, Susan
; APPLICANT: Singh, Manohan
; APPLICANT: Ulmer, Jeffrey
; APPLICANT: Dubensky, Jr., Thomas W.
; TITLE OF INVENTION: MICROPARTICLES FOR DELIVERY OF HETEROLOGOUS NUCLEIC ACIDS
; FILE REFERENCE: PP16269.004
; CURRENT APPLICATION NUMBER: US/09/967,464
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 60/236,105
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 60/315,905
; PRIOR FILING DATE: 2001-08-30
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 67
; LENGTH: 1479
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
US-09-967-464-67

```

```

Query Match          100.0%; Score 60; DB 10; Length 1479;
Best Local Similarity 100.0%; Pred. No. 2.6e-12;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 GACATCAAGCAGGCGCCCAAGAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 60
        |||
        844 GACATCAAGCAGGCGCCCAAGAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 903
DB

```

#### RESULT 5

```

US-09-899-575-3
; Sequence 3, Application US/09899575
; Publication No. US20030223961A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Zur Megede, Jan
; APPLICANT: Barnett, Susan W.
; APPLICANT: Egnelbrecht, Susan
; APPLICANT: van Rensburg, Estrelita Janse
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP01631.102
; CURRENT APPLICATION NUMBER: US/09/899,575
; CURRENT FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 09/475,704
; PRIOR FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1479
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic Gag
; OTHER INFORMATION: of HIV strain AF110965
US-09-899-575-3

```

```

Query Match          100.0%; Score 60; DB 13; Length 1479;
Best Local Similarity 100.0%; Pred. No. 2.6e-12;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 GACATCAAGCAGGCGCCCAAGAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 60
        |||
        844 GACATCAAGCAGGCGCCCAAGAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 903
DB

```

RESULT 6  
US-09-899-575-20

```
; Sequence 20, Application US/09899575
; Publication No. US20030223961A1
; GENERAL INFORMATION:
; APPLICANT: Zur Megede, Jan
; APPLICANT: Barnett, Susan W.
; APPLICANT: Egnelbrecht, Susan
; APPLICANT: van Rensburg, Estrelita Janse
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP01631.102
; CURRENT APPLICATION NUMBER: US/09/899,575
; CURRENT FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 09/475,704
; PRIOR FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 1479
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic Gag
; US-09-899-575-20

Query Match          100.0%; Score 60; DB 13; Length 1479;
Best Local Similarity 100.0%; Pred. No. 2.6e-12;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GACATCAAGCAGGGCCCCCAAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 60
      |||
DB      844 GACATCAAGCAGGGCCCCCAAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 903

RESULT 7
US-09-899-575-99
; Sequence 99, Application US/09899575
; Publication No. US20030223961A1
; GENERAL INFORMATION:
; APPLICANT: Zur Megede, Jan
; APPLICANT: Barnett, Susan W.
; APPLICANT: Egnelbrecht, Susan
; APPLICANT: van Rensburg, Estrelita Janse
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP01631.102
; CURRENT APPLICATION NUMBER: US/09/899,575
; CURRENT FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 09/475,704
; PRIOR FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 99
; LENGTH: 1491
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Gag_TV2_C_ZAopt
; US-09-899-575-99

Query Match          100.0%; Score 60; DB 13; Length 1491;
Best Local Similarity 100.0%; Pred. No. 2.6e-12;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GACATCAAGCAGGGCCCCCAAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 60
      |||
DB      841 GACATCAAGCAGGGCCCCCAAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 900

RESULT 8
US-09-899-575-51
; Sequence 51, Application US/09899575
; Publication No. US20030223961A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Zur Megede, Jan
; APPLICANT: Barnett, Susan W.
; APPLICANT: Egnelbrecht, Susan
; APPLICANT: van Rensburg, Estrelita Janse
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP01631.102
; CURRENT APPLICATION NUMBER: US/09/899,575
; CURRENT FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 09/475,704
; PRIOR FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 51
; LENGTH: 1494
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: HIV Type C
; US-09-899-575-51

Query Match          100.0%; Score 60; DB 13; Length 1494;
Best Local Similarity 100.0%; Pred. No. 2.6e-12;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GACATCAAGCAGGGCCCCCAAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 60
      |||
DB      841 GACATCAAGCAGGGCCCCCAAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 900

RESULT 9
US-10-190-435-20
; Sequence 20, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 2742
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GagTatRevNef_C
; US-10-190-435-20

Query Match          100.0%; Score 60; DB 15; Length 2742;
Best Local Similarity 100.0%; Pred. No. 2.7e-12;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GACATCAAGCAGGGCCCCCAAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 60
      |||
DB      850 GACATCAAGCAGGGCCCCCAAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 909

RESULT 10
US-10-190-435-57
; Sequence 57, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
```



```
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 57
; LENGTH: 2742
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: TatRevNefgag C
;
US-10-190-435-57
```

```
Query Match      100.0%; Score 60; DB 15; Length 2742;
Best Local Similarity 100.0%; Pred. No. 2.7e-12;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 GACATCAAGCAGGCGCCCAAGAGCCCTTCGCGACTACGTGACCGCTTCTCAAGACC 60
      |||
Db      2104 GACATCAAGCAGGCGCCCAAGAGCCCTTCGCGACTACGTGACCGCTTCTCAAGACC 2163
```

## RESULT 11

```
US-10-190-305A-15
; Sequence 15, Application US/10190305A
; Publication No. US20030198621A1
; GENERAL INFORMATION:
```

```
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan
; APPLICANT: LIAN, Ying
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B AND/OR
; TITLE OF INVENTION: TYPE C POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 2302-18702 / 18702.002
; CURRENT APPLICATION NUMBER: US/10/190,305A
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 2742
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GagTatRevNef_C
;
US-10-190-305A-15
```

```
Query Match      100.0%; Score 60; DB 15; Length 2742;
Best Local Similarity 100.0%; Pred. No. 2.7e-12;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 GACATCAAGCAGGCGCCCAAGAGCCCTTCGCGACTACGTGACCGCTTCTCAAGACC 60
      |||
Db      850 GACATCAAGCAGGCGCCCAAGAGCCCTTCGCGACTACGTGACCGCTTCTCAAGACC 909
```

## RESULT 12

```
US-10-190-305A-81
; Sequence 81, Application US/10190305A
; Publication No. US20030198621A1
; GENERAL INFORMATION:
```

```
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan
; APPLICANT: LIAN, Ying
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B AND/OR
; TITLE OF INVENTION: TYPE C POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 2302-18702 / 18702.002
; CURRENT APPLICATION NUMBER: US/10/190,305A
; NUMBER OF SEQ ID NOS: 93
```

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 81
; LENGTH: 2742
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: TatRevNefgag C
;
US-10-190-305A-81
```

```
Query Match      100.0%; Score 60; DB 15; Length 2742;
Best Local Similarity 100.0%; Pred. No. 2.7e-12;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 GACATCAAGCAGGCGCCCAAGAGCCCTTCGCGACTACGTGACCGCTTCTCAAGACC 60
      |||
Db      2104 GACATCAAGCAGGCGCCCAAGAGCCCTTCGCGACTACGTGACCGCTTCTCAAGACC 2163
```

## RESULT 13

```
US-10-190-435-18
; Sequence 18, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
```

```
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 3162
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GagRTmut_C
;
US-10-190-435-18
```

```
Query Match      100.0%; Score 60; DB 15; Length 3162;
Best Local Similarity 100.0%; Pred. No. 2.7e-12;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 GACATCAAGCAGGCGCCCAAGAGCCCTTCGCGACTACGTGACCGCTTCTCAAGACC 60
      |||
Db      850 GACATCAAGCAGGCGCCCAAGAGCCCTTCGCGACTACGTGACCGCTTCTCAAGACC 909
```

## RESULT 14

```
US-10-190-435-16
; Sequence 16, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
```

```
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 3462
; TYPE: DNA
; ORGANISM: Artificial Sequence
```

FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: GagProteinRTmut\_C  
US-10-190-435-16

Query Match 100.0%; Score 60; DB 15; Length 3462;  
Best Local Similarity 100.0%; Pred. No. 2.7e-12;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATCAAGCAGCGGGCCCAAGAGCCCTTCGCGACTACGTGAGCCGCTTCTTCAAGACC 60  
|||  
DB 850 GACATCAAGCAGCGGGCCCAAGAGCCCTTCGCGACTACGTGAGCCGCTTCTTCAAGACC 909

RESULT 15  
US-10-190-435-13

Sequence 13, Application US/10190435  
Publication No. US20030143248A1  
GENERAL INFORMATION:  
APPLICANT: ZUR MEGEDE, Jan  
APPLICANT: BARNETT, Susan W.  
APPLICANT: LIAN, Ying  
APPLICANT: ENGELBRECHT, Susan  
APPLICANT: VAN RENSBURG, Estrelita J.  
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C  
TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF  
FILE REFERENCE: P18133.003 / 2302-18133  
CURRENT APPLICATION NUMBER: US/10/190,435  
CURRENT FILING DATE: 2002-12-30  
NUMBER OF SEQ ID NOS: 319  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 13  
LENGTH: 3531  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: GagPolmut\_C  
US-10-190-435-13

Query Match 100.0%; Score 60; DB 15; Length 3531;  
Best Local Similarity 100.0%; Pred. No. 2.7e-12;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATCAAGCAGCGGGCCCAAGAGCCCTTCGCGACTACGTGAGCCGCTTCTTCAAGACC 60  
|||  
DB 850 GACATCAAGCAGCGGGCCCAAGAGCCCTTCGCGACTACGTGAGCCGCTTCTTCAAGACC 909

Search completed: May 28, 2004, 16:09:15  
Job time : 31.7692 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: May 28, 2004, 09:05:41 ; Search time 246.567 Seconds  
(without alignments)  
10547.179 Million cell updates/sec

Title: US-09-475-704A-1  
Perfect score: 60  
Sequence: 1 gacatcaagcaggccccca.....tggaccgtcttctcaagacc 60

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 2167151695 residues  
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenBmb1:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vi:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pln:\*  
35: em\_htg\_rod:\*  
36: em\_htg\_mam:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	60	100.0	60	6	AX455885	AX455885 Sequence
2	60	100.0	60	6	AX455937	AX455937 Sequence
3	60	100.0	1479	6	AX455887	AX455887 Sequence
4	60	100.0	1479	6	AX455904	AX455904 Sequence
5	60	100.0	1479	6	AX468543	AX468543 Sequence
6	60	100.0	1479	6	AX468547	AX468547 Sequence
7	60	100.0	1485	12	AY181195	AY181195 Synthetic
8	60	100.0	1491	6	AX455983	AX455983 Sequence
9	60	100.0	1494	6	AX455935	AX455935 Sequence
10	60	100.0	9166	6	AX427930	AX427930 Sequence
11	60	100.0	9169	6	AX427931	AX427931 Sequence
12	60	100.0	9785	6	AX427938	AX427938 Sequence
13	60	100.0	9788	6	AX427936	AX427936 Sequence
14	55.2	92.0	60	6	AX455886	AX455886 Sequence
15	55.2	92.0	1509	6	AX455888	AX455888 Sequence
16	55.2	92.0	1509	6	AX455905	AX455905 Sequence
17	55.2	92.0	1509	6	AX468544	AX468544 Sequence
18	55.2	92.0	1509	6	AX468548	AX468548 Sequence
19	55.2	92.0	1608	6	AX188558	AX188558 Sequence
20	55.2	92.0	1914	6	AX188560	AX188560 Sequence
21	55.2	92.0	2493	6	AX188562	AX188562 Sequence
22	55.2	92.0	4350	6	AX188564	AX188564 Sequence
23	53.6	89.3	60	6	BD263651	BD263651 Improved
24	53.6	89.3	60	6	AR373334	AR373334 Sequence
25	53.6	89.3	1268	6	BD263642	BD263642 Improved
26	53.6	89.3	1268	6	AR373325	AR373325 Sequence
27	53.6	89.3	1509	12	AF201927	AF201927 Synthetic
28	53.6	89.3	1515	6	BD263637	BD263637 Improved
29	53.6	89.3	1515	6	AR373320	AR373320 Sequence
30	53.6	89.3	1847	12	AF202464	AF202464 Synthetic
31	53.6	89.3	1847	12	AF202465	AF202465 Synthetic
32	53.6	89.3	1853	6	BD263638	BD263638 Improved
33	53.6	89.3	1853	6	AR373321	AR373321 Sequence
34	53.6	89.3	1865	6	BD263700	BD263700 Improved
35	53.6	89.3	1865	6	BD263701	BD263701 Improved
36	53.6	89.3	1865	6	AR373383	AR373383 Sequence
37	53.6	89.3	1865	6	AR373384	AR373384 Sequence
38	53.6	89.3	2031	6	BD263640	BD263640 Improved
39	53.6	89.3	2031	6	AR373323	AR373323 Sequence
40	53.6	89.3	4319	6	BD263639	BD263639 Improved
41	53.6	89.3	4319	6	AR373322	AR373322 Sequence
42	53.6	89.3	4472	6	BD263697	BD263697 Improved
43	53.6	89.3	4472	6	AR373380	AR373380 Sequence
44	53.6	89.3	4608	6	BD263698	BD263698 Improved
45	53.6	89.3	4608	6	AR373381	AR373381 Sequence

#### ALIGNMENTS

RESULT 1  
AX455885  
LOCUS AX455885 60 bp DNA linear PAT 06-JUL-2002  
DEFINITION Sequence 1 from Patent WO0204493.  
ACCESSION AX455885  
VERSION AX455885.1 GI:21714878  
KEYWORDS  
SOURCE Human immunodeficiency virus  
ORGANISM Human immunodeficiency virus  
Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate  
lentivirus group.  
REFERENCE  
AUTHORS zur Megede, J., Barnett, S.W., Engelbrecht, S. and van Rensburg, E.  
TITLE Polynucleotides encoding antigenic hiv type c polypeptides,  
polypeptides and uses thereof

JOURNAL Patent: WO 0204493-A 1 17-JAN-2002;  
CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)  
FEATURES  
source  
1. .60  
/organism="Human immunodeficiency virus"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:12721"  
ORIGIN  
Query Match 100.0%; Score 60; DB 6; Length 60;  
Best Local Similarity 100.0%; Pred. No. 2.8e-07;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GACATCAAGCAGGGCCCCCAAGAGACCCCTTCGCGACTACGTGACCGCTTCTTCAAGACC 60  
1 GACATCAAGCAGGGCCCCCAAGAGACCCCTTCGCGACTACGTGACCGCTTCTTCAAGACC 60  
Db 1 GACATCAAGCAGGGCCCCCAAGAGACCCCTTCGCGACTACGTGACCGCTTCTTCAAGACC 60  
RESULT 2  
AX455937 60 bp DNA linear PAT 06-JUL-2002  
LOCUS  
DEFINITION Sequence 53 from Patent WO0204493.  
ACCESSION AX455937  
VERSION AX455937.1 GI:21714921  
KEYWORDS  
SOURCE  
ORGANISM  
synthetic construct  
synthetic construct  
artificial sequences.  
REFERENCE  
AUTHORS  
TITL  
1 zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,B.  
Polynucleotides encoding antigenic hiv type c polypeptides,  
polypeptides and uses thereof  
Patent: WO 0204493-A 53 17-JAN-2002;  
CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)  
FEATURES  
source  
1. .60  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="HIV Type C Gag Major Homology Region Optimized"  
ORIGIN  
Query Match 100.0%; Score 60; DB 6; Length 60;  
Best Local Similarity 100.0%; Pred. No. 2.8e-07;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GACATCAAGCAGGGCCCCCAAGAGACCCCTTCGCGACTACGTGACCGCTTCTTCAAGACC 60  
1 GACATCAAGCAGGGCCCCCAAGAGACCCCTTCGCGACTACGTGACCGCTTCTTCAAGACC 60  
Db 1 GACATCAAGCAGGGCCCCCAAGAGACCCCTTCGCGACTACGTGACCGCTTCTTCAAGACC 60  
RESULT 3  
AX455887 1479 bp DNA linear PAT 06-JUL-2002  
LOCUS  
DEFINITION Sequence 3 from Patent WO0204493.  
ACCESSION AX455887  
VERSION AX455887.1 GI:21714880  
KEYWORDS  
SOURCE  
ORGANISM  
synthetic construct  
synthetic construct  
artificial sequences.  
REFERENCE  
AUTHORS  
TITL  
1 zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,B.  
Polynucleotides encoding antigenic hiv type c polypeptides,  
polypeptides and uses thereof  
Patent: WO 0204493-A 3 17-JAN-2002;  
CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)  
FEATURES  
source  
1. .1479  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="synthetic Gag of HIV strain AF110965"

ORIGIN  
Query Match 100.0%; Score 60; DB 6; Length 1479;  
Best Local Similarity 100.0%; Pred. No. 1.9e-07;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GACATCAAGCAGGGCCCCCAAGAGACCCCTTCGCGACTACGTGACCGCTTCTTCAAGACC 60  
1 GACATCAAGCAGGGCCCCCAAGAGACCCCTTCGCGACTACGTGACCGCTTCTTCAAGACC 60  
Db 844 GACATCAAGCAGGGCCCCCAAGAGACCCCTTCGCGACTACGTGACCGCTTCTTCAAGACC 903  
RESULT 4  
AX455904 1479 bp DNA linear PAT 06-JUL-2002  
LOCUS  
DEFINITION Sequence 20 from Patent WO0204493.  
ACCESSION AX455904  
VERSION AX455904.1 GI:21714896  
KEYWORDS  
SOURCE  
ORGANISM  
synthetic construct  
synthetic construct  
artificial sequences.  
REFERENCE  
AUTHORS  
TITL  
1 zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,B.  
Polynucleotides encoding antigenic hiv type c polypeptides,  
polypeptides and uses thereof  
Patent: WO 0204493-A 20 17-JAN-2002;  
CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)  
FEATURES  
source  
1. .1479  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="synthetic Gag coding sequence of HIV strain AF110965"  
ORIGIN  
Query Match 100.0%; Score 60; DB 6; Length 1479;  
Best Local Similarity 100.0%; Pred. No. 1.9e-07;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GACATCAAGCAGGGCCCCCAAGAGACCCCTTCGCGACTACGTGACCGCTTCTTCAAGACC 60  
1 GACATCAAGCAGGGCCCCCAAGAGACCCCTTCGCGACTACGTGACCGCTTCTTCAAGACC 60  
Db 844 GACATCAAGCAGGGCCCCCAAGAGACCCCTTCGCGACTACGTGACCGCTTCTTCAAGACC 903  
RESULT 5  
AX468543 1479 bp DNA linear PAT 16-JUL-2002  
LOCUS  
DEFINITION Sequence 63 from Patent WO0226209.  
ACCESSION AX468543  
VERSION AX468543.1 GI:21901373  
KEYWORDS  
SOURCE  
ORGANISM  
Human immunodeficiency virus 1 (HIV-1)  
Human immunodeficiency virus 1  
viruses; Retroid viruses; Retroviridae; Lentivirus; Primate  
lentivirus group.  
REFERENCE  
AUTHORS  
TITL  
1 O'Hagan,D., Otten,G., Donnelly,J.J., Polo,J.M., Barnett,S.,  
Singh,M., Ulmer,J. and Dubensky,T.W.  
Microparticles for delivery of the heterologous nucleic acids  
Patent: WO 0226209-A 63 04-APR-2002;  
CHIRON CORPORATION (US)  
FEATURES  
source  
1. .1479  
/organism="Human immunodeficiency virus 1"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:11676"  
ORIGIN  
Query Match 100.0%; Score 60; DB 6; Length 1479;  
Best Local Similarity 100.0%; Pred. No. 1.9e-07;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACATCAAGCAGGGCCCCCAAGAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 60  
|||||  
Db 844 GACATCAAGCAGGGCCCCCAAGAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 903

RESULT 6  
AX468547 1479 bp DNA linear PAT 16-JUL-2002  
DEFINITION Sequence 67 from Patent WO0226209.  
ACCESSION AX468547  
VERSION AX468547.1 GI:21901377  
KEYWORDS  
SOURCE Human immunodeficiency virus 1 (HIV-1)  
ORGANISM Human immunodeficiency virus 1  
viruses; Retrovird viruses; Retroviridae; Lentivirus; Primate  
lentivirus group.

REFERENCE  
1 O'Hagan,D., Otten,G., Donnelly,J.J., Polo,J.M., Barnett,S., Singh,M., Ulmer,J. and Dubensky,T.M.  
TITLE Microparticles for delivery of the heterologous nucleic acids  
JOURNAL Patent: WO 0226209-A 67 04-APR-2002;  
CHIRON CORPORATION (US)  
FEATURES  
source Location/Qualifiers  
1. 1479  
/organism="Human immunodeficiency virus 1"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:11676"

ORIGIN  
Query Match 100.0%; Score 60; DB 6; Length 1479;  
Best Local Similarity 100.0%; Pred. No. 1.9e-07;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACATCAAGCAGGGCCCCCAAGAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 60  
|||||  
Db 844 GACATCAAGCAGGGCCCCCAAGAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 903

RESULT 7  
AY181195 1485 bp DNA linear SYN 02-OCT-2003  
LOCUS AY181195  
DEFINITION Synthetic construct HIV-1-derived gag protein (gag) gene, complete  
cgs.  
ACCESSION AY181195  
VERSION AY181195.1 GI:37413992  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
artificial sequences.  
REFERENCE  
1 (bases 1 to 1485)  
Gao,F., Li,Y., Decker,J.M., Peyerl,F.W., Bibollet-Ruche,F., Rodenburg,C.M., Chen,Y., Shaw,D.R., Allen,S., Musonda,R., Shaw,G.M., Zajac,A.J., Letvin,N. and Hahn,B.H.  
TITLE Codon usage optimization of HIV type 1 subtype C gag, pol, env, and  
nef genes: in vitro expression and immune responses in  
DNA-vaccinated mice  
JOURNAL AIDS Res. Hum. Retroviruses 19 (9), 817-823 (2003)  
AUTHORS  
2 (bases 1 to 1485)  
Gao,F., Li,Y., Decker,J.M., Bibollet-Ruche,F., Rodenburg,C.M., Chen,Y., Shaw,D.R., Allen,S., Musonda,R., Shaw,G.M., Zajac,A.J. and Hahn,B.H.  
TITLE Direct Submission  
JOURNAL Submitted (16-NOV-2002) Medicine, Duke University Medical Center,  
112 RPIII, LaSalle Street, Durham, NC 27710, USA  
FEATURES  
source Location/Qualifiers  
1. 1485  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
/note="codon usage optimization of HIV-1 96ZM651"  
1. 1485  
/gene="gag"  
1. 1485  
CDS

/gene="gag"  
/codon\_start=1  
/transl\_table=11  
/product="gag protein"  
/protein\_id="AA065393.1"  
/db\_xref="GI:37413993"  
/translation="MGARASILRGKLDKWEKIRLRPGKRRYMIKHLVWASRELERF  
ALNGLLETSEGGQIMKQLPALQTEBELSLVNTVATLYCVHEGVEVRDTKALD  
RIEERQNKIQOKIQKIQKIQKQADGKVSQNPVQNLQGMVHOKLSPTLNAAWVKYTE  
KAPSPVETPMFTALSGATPDQDNLNTVGGHQAQMLKDTINEEAEMDRLLHPVH  
AGPIAPGQMRBPRGSDIAGTSTLQROIAMTNSNPPIVGDYKRWIILGLNKIVRM  
SPVSLIDIKQGPKEPRDYVDRPFKTLPAEQATQEVKMMTDTLLVQNAIPDCKTILK  
ALGPATLEBMTACQGVGSPSHKARVLAEMSQTNSVNIIMQSNFKGNKRMVCFN  
CGKEGHIARNCRAPRRKCGMKCGKHQMKDCTERQANFLGKIWP SHKGRPGNFLQNR  
PEPTADPAESFRFEETTPAPKQESKDRREALTSLKSLFGSDPLSQ"

ORIGIN  
Query Match 100.0%; Score 60; DB 12; Length 1485;  
Best Local Similarity 100.0%; Pred. No. 1.9e-07;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACATCAAGCAGGGCCCCCAAGAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 60  
|||||  
Db 847 GACATCAAGCAGGGCCCCCAAGAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 906

RESULT 8  
AX455983 1491 bp DNA linear PAT 06-JUL-2002  
LOCUS AX455983  
DEFINITION Sequence 99 from Patent WO0204493.  
ACCESSION AX455983  
VERSION AX455983.1 GI:21714967  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
artificial sequences.  
REFERENCE  
1 zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,B.  
AUTHORS  
TITLE Polynucleotides encoding antigenic hiv type c polypeptides,  
polypeptides and uses thereof  
JOURNAL Patent: WO 0204493-A 99 17-JAN-2002;  
CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)  
FEATURES  
source Location/Qualifiers  
1. 1491  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="Gag\_TV2\_C\_ZAopt"

ORIGIN  
Query Match 100.0%; Score 60; DB 6; Length 1491;  
Best Local Similarity 100.0%; Pred. No. 1.9e-07;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACATCAAGCAGGGCCCCCAAGAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 60  
|||||  
Db 841 GACATCAAGCAGGGCCCCCAAGAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 900

RESULT 9  
AX455935 1494 bp DNA linear PAT 06-JUL-2002  
LOCUS AX455935  
DEFINITION Sequence 51 from Patent WO0204493.  
ACCESSION AX455935  
VERSION AX455935.1 GI:21714919  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
artificial sequences.  
REFERENCE  
1 zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,B.  
AUTHORS  
TITLE Polynucleotides encoding antigenic hiv type c polypeptides,  
polypeptides and uses thereof



JOURNAL Patent: WO 0204493-A 51 17-JAN-2002;  
CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)  
FEATURES  
Source  
1. 1494  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="HIV Type C Gag optimized"  
ORIGIN  
Query Match 100.0%; Score 60; DB 6; Length 1494;  
Best Local Similarity 100.0%; Pred. No. 1.9e-07;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GACATCAAGCAGGGCCCCAAGAGAGCCCTTCGCGACTACGTGACCGCTTCTTCAAGACC 60  
|||||  
Db 841 GACATCAAGCAGGGCCCCAAGAGAGCCCTTCGCGACTACGTGACCGCTTCTTCAAGACC 900  
RESULT 10  
LOCUS AX427930 9166 bp DNA linear PAT 20-JUN-2002  
DEFINITION Sequence 168 from Patent WO0232943.  
ACCESSION AX427930  
VERSION AX427930.1 GI:21538017  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
FEATURES  
1  
AUTHORS Huang, Y. and Nabel, G.J.  
TITLE Modifications of hiv env, gag, and pol enhance immunogenicity for  
genetic immunization  
JOURNAL Patent: WO 0232943-A 168 25-APR-2002;  
GOVERNMENT OF THE UNITED STATES (US)  
ORIGIN  
Location/Qualifiers  
1. 9166  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="plasmid pVR1012x/s containing HIV genes"  
ORIGIN  
Query Match 100.0%; Score 60; DB 6; Length 9166;  
Best Local Similarity 100.0%; Pred. No. 1.5e-07;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GACATCAAGCAGGGCCCCAAGAGAGCCCTTCGCGACTACGTGACCGCTTCTTCAAGACC 60  
|||||  
Db 2723 GACATCAAGCAGGGCCCCAAGAGAGCCCTTCGCGACTACGTGACCGCTTCTTCAAGACC 2782  
RESULT 11  
LOCUS AX427931 9169 bp DNA linear PAT 20-JUN-2002  
DEFINITION Sequence 169 from Patent WO0232943.  
ACCESSION AX427931  
VERSION AX427931.1 GI:21538018  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
FEATURES  
1  
AUTHORS Huang, Y. and Nabel, G.J.  
TITLE Modifications of hiv env, gag, and pol enhance immunogenicity for  
genetic immunization  
JOURNAL Patent: WO 0232943-A 169 25-APR-2002;  
GOVERNMENT OF THE UNITED STATES (US)  
ORIGIN  
Location/Qualifiers  
1. 9169  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"

/note="plasmid pVR1012x/s containing HIV genes"  
ORIGIN  
Query Match 100.0%; Score 60; DB 6; Length 9169;  
Best Local Similarity 100.0%; Pred. No. 1.5e-07;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GACATCAAGCAGGGCCCCAAGAGAGCCCTTCGCGACTACGTGACCGCTTCTTCAAGACC 60  
|||||  
Db 2729 GACATCAAGCAGGGCCCCAAGAGAGCCCTTCGCGACTACGTGACCGCTTCTTCAAGACC 2788  
RESULT 12  
LOCUS AX427938 9785 bp DNA linear PAT 20-JUN-2002  
DEFINITION Sequence 176 from Patent WO0232943.  
ACCESSION AX427938  
VERSION AX427938.1 GI:21538025  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
FEATURES  
1  
AUTHORS Huang, Y. and Nabel, G.J.  
TITLE Modifications of hiv env, gag, and pol enhance immunogenicity for  
genetic immunization  
JOURNAL Patent: WO 0232943-A 176 25-APR-2002;  
GOVERNMENT OF THE UNITED STATES (US)  
ORIGIN  
Location/Qualifiers  
1. 9785  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="plasmid pVR1012x/s containing HIV genes"  
ORIGIN  
Query Match 100.0%; Score 60; DB 6; Length 9785;  
Best Local Similarity 100.0%; Pred. No. 1.5e-07;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GACATCAAGCAGGGCCCCAAGAGAGCCCTTCGCGACTACGTGACCGCTTCTTCAAGACC 60  
|||||  
Db 2729 GACATCAAGCAGGGCCCCAAGAGAGCCCTTCGCGACTACGTGACCGCTTCTTCAAGACC 2788  
RESULT 13  
LOCUS AX427936 9788 bp DNA linear PAT 20-JUN-2002  
DEFINITION Sequence 174 from Patent WO0232943.  
ACCESSION AX427936  
VERSION AX427936.1 GI:21538023  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
FEATURES  
1  
AUTHORS Huang, Y. and Nabel, G.J.  
TITLE Modifications of hiv env, gag, and pol enhance immunogenicity for  
genetic immunization  
JOURNAL Patent: WO 0232943-A 174 25-APR-2002;  
GOVERNMENT OF THE UNITED STATES (US)  
ORIGIN  
Location/Qualifiers  
1. 9788  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="plasmid pVR1012x/s containing HIV genes"  
ORIGIN  
Query Match 100.0%; Score 60; DB 6; Length 9788;  
Best Local Similarity 100.0%; Pred. No. 1.5e-07;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATCAAGCAGGGCCCCCAAGAGCCCTTCCGCGACTAGTGAACCGCTTCTCAAGACC 60  
|||||  
Db 2723 GACATCAAGCAGGGCCCCCAAGAGCCCTTCCGCGACTAGTGAACCGCTTCTCAAGACC 2782

## RESULT 14

LOCUS AX455886 60 bp DNA linear PAT 06-JUL-2002  
DEFINITION Sequence 2 from Patent WO0204493.  
ACCESSION AX455886  
VERSION AX455886.1 GI:21714879  
KEYWORDS  
SOURCE Human immunodeficiency virus  
ORGANISM Human immunodeficiency virus  
Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate  
lentivirus group.

REFERENCE 1  
AUTHORS zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.  
TITLE Polynucleotides encoding antigenic hiv type c polypeptides,  
polypeptides and uses thereof  
JOURNAL Patent: WO 0204493-A 2 17-JAN-2002;  
CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)

FEATURES  
source 1..60  
/organism="Human immunodeficiency virus"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:12721"

## ORIGIN

Query Match 92.0%; Score 55.2; DB 6; Length 60;  
Best Local Similarity 95.0%; Pred.No. 7.2e-06;  
Matches 57; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GACATCAAGCAGGGCCCCCAAGAGCCCTTCCGCGACTAGTGAACCGCTTCTCAAGACC 60  
|||||  
Db 1 GACATCCGCCAGGGCCCCCAAGAGCCCTTCCGCGACTAGTGAACCGCTTCTCAAGACC 60

## RESULT 15

LOCUS AX455888 1509 bp DNA linear PAT 06-JUL-2002  
DEFINITION Sequence 4 from Patent WO0204493.  
ACCESSION AX455888  
VERSION AX455888.1 GI:21714881  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
artificial sequences.

REFERENCE 1  
AUTHORS zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.  
TITLE Polynucleotides encoding antigenic hiv type c polypeptides,  
polypeptides and uses thereof  
JOURNAL Patent: WO 0204493-A 4 17-JAN-2002;  
CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)

FEATURES  
source 1..1509  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="synthetic Gag of HIV strain AF110967"

## ORIGIN

Query Match 92.0%; Score 55.2; DB 6; Length 1509;  
Best Local Similarity 95.0%; Pred.No. 4.8e-06;  
Matches 57; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GACATCAAGCAGGGCCCCCAAGAGCCCTTCCGCGACTAGTGAACCGCTTCTCAAGACC 60  
|||||  
Db 841 GACATCCGCCAGGGCCCCCAAGAGCCCTTCCGCGACTAGTGAACCGCTTCTCAAGACC 900

Search completed: May 28, 2004, 13:23:15  
Job time : 247.567 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 28, 2004, 07:10:24 ; Search time 27.8049 Seconds  
(without alignments)  
9167.164 Million cell updates/sec

Title: US-09-475-704A-1

Perfect score: 60  
Sequence: 1 gacatcaagcaggcgcccaaa.....tggacgcgtctctcaagacc 60

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_29Jan04:\*

1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002s:\*  
7: geneseqn2003as:\*  
8: geneseqn2003bs:\*  
9: geneseqn2003cs:\*  
10: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	3	AAA51607 HIV Gag m
2	60	100.0	60	6	ABL39974 Synthetic
3	60	100.0	1479	3	AAA51609 HIV synth
4	60	100.0	1479	3	AAA51625 HIV codon
5	60	100.0	1479	6	AA144548 HIV-1 p55
6	60	100.0	1479	6	AA144552 HIV-1 p55
7	60	100.0	1479	6	ABL39953 Synthetic
8	60	100.0	1479	6	ABL39957 Synthetic
9	60	100.0	1491	6	ABL40020 Synthetic
10	60	100.0	1494	6	ABL39972 Synthetic
11	60	100.0	2742	7	ACA03590 Synthetic
12	60	100.0	2742	7	ACA03524 Synthetic
13	60	100.0	2742	9	ADC13278 DNA of HI
14	60	100.0	2742	9	ADC13241 DNA of HI
15	60	100.0	3162	9	ADC13239 DNA of HI
16	60	100.0	3462	9	ADC13237 DNA of HI
17	60	100.0	3531	9	ADC13234 DNA of HI
18	60	100.0	3537	9	ADC13236 DNA of HI
19	60	100.0	3538	9	ADC13235 DNA of HI
20	60	100.0	3930	9	ADC13230 DNA of HI
21	60	100.0	3930	9	ADC13231 DNA of HI
22	60	100.0	3930	9	ADC13232 DNA of HI
23	60	100.0	4419	7	ACA03523 Synthetic

24	60	100.0	4419	9	ADC13240 DNA of HI
25	60	100.0	4423	9	ADC13256 DNA of HI
26	60	100.0	4546	9	ADC13255 DNA of HI
27	60	100.0	4615	9	ADC13257 DNA of HI
28	60	100.0	4702	9	ADC13259 DNA of HI
29	60	100.0	4713	9	ACA03592 Synthetic
30	60	100.0	4713	7	ADC13280 DNA of HI
31	60	100.0	4716	7	ACA03522 Synthetic
32	60	100.0	4716	9	ADC13238 DNA of HI
33	60	100.0	5145	9	ACA03521 Synthetic
34	60	100.0	5145	9	ADC13233 DNA of HI
35	60	100.0	5184	7	ACA03591 Synthetic
36	60	100.0	5184	9	ADC13279 DNA of HI
37	60	100.0	9166	6	ABK91616 Modified
38	60	100.0	9169	6	ABK91617 Modified
39	60	100.0	9785	6	ABK91624 Modified
40	60	100.0	9788	6	ABK91622 Modified
41	58.4	97.3	1092	7	AAD53096 Gag p17/2
42	58.4	97.3	1092	7	AAD53097 Gag p17/2
43	58.4	97.3	3020	7	Aad53105 Human imm
44	58.4	97.3	3021	7	AAD53106 Human imm
45	58.4	97.3	3021	7	AAD53107 Human imm

ALIGNMENTS

RESULT 1  
ID AAA51607 standard; DNA; 60 BP.  
XX AAA51607;  
AC  
XX  
DT 31-OCT-2000 (first entry)  
XX  
XX HIV Gag major homology region nucleotides 844-903.  
DE  
XX  
KW Gag; expression cassette; antigenic; type C; HIV; Env; synthetic;  
KW DNA immunization; packaging cell line; antigen presentation; ss.  
XX  
OS Human immunodeficiency virus.  
XX  
PN WO200039304-A2.  
XX  
PD 06-JUL-2000.  
XX  
PF 30-DEC-1999; 99WO-US031273.  
XX  
PR 31-DEC-1998; 98US-0114495P.  
PR 01-SEP-1999; 99US-0152195P.  
XX  
PA (CHIR ) CHIRON CORP.  
XX  
PI Barnett S, Zur Megede J;  
XX  
PI WPI; 2000-452401/39.  
XX  
PT Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV Env  
PT polypeptide and the polypeptide useful for immunizing a mammal especially  
PT human against HIV.  
XX  
PS Claim 1; Page 92; 113pp; English.  
XX  
CC Expression cassettes comprising a polynucleotide encoding antigenic type  
CC C human immunodeficiency virus (HIV) Gag or Env polypeptides are useful  
CC in DNA immunization, generation of packaging cell lines and production of  
CC Gag- and/or Env-containing proteins. Synthetic Env and Gag expression  
CC cassettes exhibit increased potency for induction of cytotoxic T-  
CC lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1 self-  
CC assemble into non-infectious virus-like particles which are used as a  
CC matrix for the proper presentation of an antigen entrapped or associated  
CC to the immune system of the host



```
XX HIV codon-optimized synthetic Gag polynucleotide.
DE
XX
XX Gag; expression cassette; antigenic; type C; HIV; Env; synthetic;
KW DNA immunization; packaging cell line; antigen presentation; ss.
XX
OS Human immunodeficiency virus; type C strain AF110965.
OS Synthetic.
XX WO200039304-A2.
XX
XX 06-JUL-2000.
XX
XX 30-DEC-1999; 99WO-US031273.
XX
XX 31-DEC-1998; 98US-011495P.
PR 01-SEP-1999; 99US-0152195P.
XX
XX (CHIR ) CHIRON CORP.
XX
XX Barnett S, Zur Megede J;
PI
XX WPI; 2000-452401/39.
XX
XX Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV Env
PT polypeptide and the polypeptide useful for immunizing a mammal especially
PT human against HIV.
XX
XX Disclosure; Page 103-104; 113pp; English.
XX
XX Expression cassettes comprising a polynucleotide encoding antigenic type
CC C human immunodeficiency virus (HIV) Gag or Env polypeptides are useful
CC in DNA immunization, generation of packaging cell lines and production of
CC Gag- and/or Env-containing proteins. Synthetic Env and Gag expression
CC cassettes exhibit increased potency for induction of cytotoxic T-
CC lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1 self-
CC assemble into non-infectious virus-like particles which are used as a
CC matrix for the proper presentation of an antigen entrapped or associated
CC to the immune system of the host
XX
SQ Sequence 1479 BP; 325 A; 533 C; 461 G; 160 T; 0 U; 0 Other;

Query Match 100.0%; Score 60; DB 3; Length 1479;
Best Local Similarity 100.0%; Pred. No. 1.3e-09;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATCAAGCAGGGCCCCCAAGAGCCCTTCGCGACTACGTGACCGCTTCTTCAAGACC 60
Db 844 GACATCAAGCAGGGCCCCCAAGAGCCCTTCGCGACTACGTGACCGCTTCTTCAAGACC 903

RESULT 5
AAL44548
ID AAL44548 standard; DNA; 1479 BP.
XX
XX AAL44548;
AC
XX 29-AUG-2003 (revised)
DT 08-NOV-2002 (first entry)
XX
XX HIV-1 p55gag polypeptide coding sequence 1.
DB
XX HIV; ds; vaccine; gene; immune response; microparticle;
KW adsorbent surface; poly(alpha-hydroxy acid); polyhydroxy butyric acid;
KW polycaprolactone; polyorthoester; polycyanoacrylate; detergent;
KW submicron emulsion; viral infection; bacterial infection;
KW parasitic infection; HIV-1 p55gag polypeptide.
XX
XX Human immunodeficiency virus 1.
OS
XX WO200226209-A2.
PN
XX 04-APR-2002.
PD
```

```
XX 28-SEP-2001; 2001WO-US030540.
PF
XX
XX 28-SEP-2000; 2000US-0236105P.
PR 30-AUG-2001; 2001US-0315905P.
XX
XX (CHIR ) CHIRON CORP.
XX
XX O'hagan D, Otten G, Donnelly JF, Polo JM, Barnett S, Singh M;
PI Ulmer J, Dubensky TW;
PI
XX WPI; 2002-519084/55.
XX
XX A microparticle to which a biologically active macromolecule is adsorbed,
PT for use as a vaccine composition to treat viral, bacterial or parasitic
PT infections, comprises a polymer microparticle, a detergent and a
PT submicron emulsion.
XX
XX Claim 72; Fig 1; 100pp; English.
XX
XX The invention relates to a method of raising an immune response in a host
CC animal. The method of the invention comprises administering a
CC microparticle that has an adsorbent surface to which a first biologically
CC active macromolecule (e.g. a nucleic acid) has been adsorbed. The
CC microparticle comprises a polymer microparticle of poly(alpha-hydroxy
CC acid), a polyhydroxy butyric acid, a polycaprolactone, a polyorthoester,
CC a polycyanoacrylate, a detergent, and submicron emulsion. The method/
CC microparticle of the invention is useful for immunising a host animal
CC against viral, bacterial or parasitic infections. The present DNA
CC sequence encodes a HIV-1 p55gag polypeptide. (Updated on 29-AUG-2003 to
CC standardise OS field)
XX
SQ Sequence 1479 BP; 325 A; 529 C; 463 G; 162 T; 0 U; 0 Other;

Query Match 100.0%; Score 60; DB 6; Length 1479;
Best Local Similarity 100.0%; Pred. No. 1.3e-09;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATCAAGCAGGGCCCCCAAGAGCCCTTCGCGACTACGTGACCGCTTCTTCAAGACC 60
Db 844 GACATCAAGCAGGGCCCCCAAGAGCCCTTCGCGACTACGTGACCGCTTCTTCAAGACC 903

RESULT 6
AAL44552
ID AAL44552 standard; DNA; 1479 BP.
XX
XX AAL44552;
AC
XX 29-AUG-2003 (revised)
DT 08-NOV-2002 (first entry)
XX
XX HIV-1 p55gag polypeptide coding sequence 3.
DB
XX HIV; ds; vaccine; gene; immune response; microparticle;
KW adsorbent surface; poly(alpha-hydroxy acid); polyhydroxy butyric acid;
KW polycaprolactone; polyorthoester; polycyanoacrylate; detergent;
KW submicron emulsion; viral infection; bacterial infection;
KW parasitic infection; HIV-1 p55gag polypeptide.
XX
XX Human immunodeficiency virus 1.
OS
XX Key Location/Qualifiers
FH variation replace(282, C)
FT variation /*tag= a
FT variation replace(552, G)
FT variation /*tag= b
FT variation replace(556, T)
FT variation /*tag= c
FT variation replace(786, G)
FT variation /*tag= d
FT variation replace(819, G)
FT variation /*tag= e
```



```
FT variation replace(1002, T)
FT /*tag= f
FT variation replace(1005, C)
FT /*tag= g
FT variation replace(1092, G)
FT /*tag= h
FT variation replace(1149, G)
FT /*tag= i
FT variation replace(1158, C)
FT /*tag= j
XX
XX WO200226209-A2.
XX
XX PD 04-APR-2002.
XX
XX PF 28-SEP-2001; 2001WO-US030540.
XX
XX PR 28-SEP-2000; 2000US-0236105P.
XX PR 30-AUG-2001; 2001US-0315905P.
XX
XX PA (CHIR ) CHIRON CORP.
XX
XX PI O'hagan D, Otten G, Donnelly JJ, Polo JM, Barnett S, Singh M,
XX PI Ulmer J, Dubensky TW;
XX
XX DR WPI; 2002-519084/55.
XX
XX CC A microparticle to which a biologically active macromolecule is adsorbed,
XX PT for use as a vaccine composition to treat viral, bacterial or parasitic
XX PT infections, comprises a polymer microparticle, a detergent and a
XX PT submicron emulsion.
XX
XX PS Claim 72; Fig 5; 100pp; English.
XX
XX CC The invention relates to a method of raising an immune response in a host
XX CC animal. The method of the invention comprises administering a
XX CC microparticle that has an adsorbent surface to which a first biologically
XX CC active macromolecule (e.g. a nucleic acid) has been adsorbed. The
XX CC microparticle comprises a polymer microparticle of poly(alpha-hydroxy
XX CC acid), a polyhydroxy butyric acid, a polycaprolactone, a polyorthoester,
XX CC a polycyanoacrylate, a detergent, and submicron emulsion. The method/
XX CC microparticle of the invention is useful for immunising a host animal
XX CC against viral, bacterial or parasitic infections. The present DNA
XX CC sequence encodes a HIV-1 p55gag polypeptide. (Updated on 29-AUG-2003 to
XX CC standardise OS field)
XX
XX SQ Sequence 1479 BP; 325 A; 534 C; 460 G; 160 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 60; DB 6; Length 1479;
XX Best Local Similarity 100.0%; Pred. No. 1.3e-09;
XX Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 GACATCAAGCAGGGCCCCAAGGAGCCCTTCGCGACTACGTGAGCCGCTTCTTCAAGACC 60
XX |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
XX DB 844 GACATCAAGCAGGGCCCCAAGGAGCCCTTCGCGACTACGTGAGCCGCTTCTTCAAGACC 903
XX
XX RESULT 7
XX ABL39953
XX ID ABL39953 standard; DNA; 1479 BP.
XX
XX AC ABL39953;
XX
XX DT 15-MAY-2002 (first entry)
XX
XX DB Synthetic Gag polynucleotide sequence SEQ ID NO:3.
XX
XX KW Human immunodeficiency virus type C; antigenic HIV type C protein;
XX KW immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
XX KW immunostimulant; gene therapy; gene; ds.
XX
XX OS Human immunodeficiency virus; type C.
XX OS Synthetic.
```

```
XX
XX PN WO200204493-A2.
XX
XX PD 17-JAN-2002.
XX
XX PF 05-JUL-2001; 2001WO-US021241.
XX
XX PR 05-JUL-2000; 2000US-00610313.
XX
XX PA (CHIR ) CHIRON CORP.
XX PA (UYST-) UNIV STELLENBOSCH.
XX
XX PI Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg EJ;
XX PI WPI; 2002-154920/20.
XX
XX DR WPI; 2002-154920/20.
XX
XX PT New polynucleotides encoding antigenic HIV Type C polypeptides, useful in
XX PT applications including DNA immunization or generation of packaging cell
XX PT lines, particularly in gene therapy.
XX
XX PS Example 1; Fig 1; 233pp; English.
XX
XX CC The present invention describes expression cassettes comprising a
XX CC polynucleotide sequence encoding a polypeptide comprising immunogenic HIV
XX CC type C polypeptides. The expression cassettes comprise any of the HIV
XX CC type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or Nef
XX CC (1). (i) have immunostimulant activity and can be used in gene therapy.
XX CC The HIV type C polynucleotides are useful in applications including DNA
XX CC immunisation, generation of packaging cell lines, and production of HIV
XX CC type C proteins. The polynucleotides are particularly useful in gene
XX CC therapy and DNA immunisation applications. ABL39942 to ABL40054 and
XX CC ABB06204 to ABB06215 represent sequences used in the exemplification of
XX CC the present invention
XX
XX SQ Sequence 1479 BP; 325 A; 529 C; 463 G; 162 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 60; DB 6; Length 1479;
XX Best Local Similarity 100.0%; Pred. No. 1.3e-09;
XX Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 GACATCAAGCAGGGCCCCAAGGAGCCCTTCGCGACTACGTGAGCCGCTTCTTCAAGACC 60
XX |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
XX DB 844 GACATCAAGCAGGGCCCCAAGGAGCCCTTCGCGACTACGTGAGCCGCTTCTTCAAGACC 903
XX
XX RESULT 8
XX ABL39957
XX ID ABL39957 standard; DNA; 1479 BP.
XX
XX AC ABL39957;
XX
XX DT 15-MAY-2002 (first entry)
XX
XX DB Synthetic Gag polynucleotide sequence SEQ ID NO:20.
XX
XX KW Human immunodeficiency virus type C; antigenic HIV type C protein;
XX KW immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
XX KW immunostimulant; gene therapy; gene; ds.
XX
XX OS Human immunodeficiency virus; type C.
XX OS Synthetic.
XX
XX PN WO200204493-A2.
XX
XX PD 17-JAN-2002.
XX
XX PF 05-JUL-2001; 2001WO-US021241.
XX
XX PR 05-JUL-2000; 2000US-00610313.
XX
XX PA (CHIR ) CHIRON CORP.
XX PA (UYST-) UNIV STELLENBOSCH.
XX
```

PI	Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg EJ;
XX	
DR	WPI; 2002-154920/20.
XX	
PT	New polynucleotides encoding antigenic HIV Type C polypeptides, useful in
PT	applications including DNA immunization or generation of packaging cell
PT	lines, particularly in gene therapy.
XX	
PS	Example 1; Fig 5; 233pp; English.
XX	
CC	The present invention describes expression cassettes comprising a
CC	polynucleotide sequence encoding a polypeptide comprising immunogenic HIV
CC	type C polypeptides. The expression cassettes comprise any of the HIV
CC	type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or Nef
CC	(I). (I) have immunostimulant activity and can be used in gene therapy.
CC	The HIV type C polynucleotides are useful in applications including DNA
CC	immunisation, generation of packaging cell lines, and production of HIV
CC	Type C proteins. The polynucleotides are particularly useful in gene
CC	therapy and DNA immunisation applications. ABL39942 to ABL40054 and
CC	ABB06204 to ABB06215 represent sequences used in the exemplification of
CC	the present invention
XX	
SEQ	Sequence 1479 BP; 325 A; 533 C; 461 G; 160 T; 0 U; 0 Other;
	Query Match 100.0%; Score 60; DB 6; Length 1479;
	Best Local Similarity 100.0%; Pred. No.1.3e-09;
	Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0
QY	1 GACATCAAGCAGGCGCCCAAGAGAGCCCTTCGCGACTACGTGAGCCGCTTCTCAAGACC 60
DB	844 GACATCAAGCAGGCGCCCAAGAGAGCCCTTCGCGACTACGTGAGCCGCTTCTCAAGACC 903
RESULT 9	
ABL40020	
ID	ABL40020 èstandard; DNA; 1491 BP.
XX	
AC	ABL40020;
XX	
DT	15-MAY-2002 (first entry)
XX	
DE	Synthetic Gag polynucleotide sequence SEQ ID NO:99.
XX	
KW	Human immunodeficiency virus type C; antigenic HIV type C protein;
KW	immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
KW	immunostimulant; gene therapy; gene; ds.
XX	
OS	Human immunodeficiency virus; type C.
OS	Synthetic.
XX	
PN	WO200204493-A2.
XX	
PD	17-JAN-2002.
XX	
PF	05-JUL-2001; 2001WO-US021241.
XX	
PR	05-JUL-2000; 2000US-00610313.
XX	
PA	(CHIR ) CHIRON CORP.
PA	(UYST-) UNIV STELLENBOSCH.
XX	
PI	Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg EJ;
XX	
DR	WPI; 2002-154920/20.
XX	
PT	New polynucleotides encoding antigenic HIV Type C polypeptides, useful in
PT	applications including DNA immunization or generation of packaging cell
PT	lines, particularly in gene therapy.
XX	
PS	Claim 8; Fig 70; 233pp; English.
XX	
CC	The present invention describes expression cassettes comprising a
CC	polynucleotide sequence encoding a polypeptide comprising immunogenic HIV

CC	type C polypeptides. The expression cassettes comprise any of the HIV
CC	type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or Nef
CC	(1). (I) have immunostimulant activity and can be used in gene therapy.
CC	The HIV type C polynucleotides are useful in applications including DNA
CC	immunisation, generation of packaging cell lines, and production of HIV
CC	Type C proteins. The polynucleotides are particularly useful in gene
CC	therapy and DNA immunisation applications. ABL39942 to ABL40054 and
CC	ABB06204 to ABB06215 represent sequences used in the exemplification of
CC	the present invention
XX	
SQ	Sequence 1491 BP; 319 A; 557 C; 457 G; 158 T; 0 U; 0 Other;
	Query Match                      100.0%; Score 60; DB 6; Length 1491;
	Best Local Similarity    100.0%; Pred. No. 1.3e-09;
	Matches    60; Conservative    0; Mismatches    0; Indels    0; Gaps    0
Dn	
Oy	1 GACATCAAGCAGGGCCCCCAAGGAGCCCTTCGCGACTACGTGACC GGCTTCTTCAAGACC 60       841 GACATCAAGCAGGGCCCCCAAGGAGCCCTTCGCGACTACGTGACC GGCTTCTTCAAGACC 900
RESULT 10	
ABL39972	
ID	ABL39972 standard; DNA; 1494 BP.
XX	
AC	ABL39972;
DT	15-MAY-2002 (first entry)
DE	Synthetic Gag polynucleotide sequence SEQ ID NO:51.
XX	
KW	Human immunodeficiency virus type C; antigenic HIV type C protein;
KM	immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
KX	immunostimulant; gene therapy; gene; ds.
OS	Human immunodeficiency virus; type C.
PN	Synthetic.
XX	
MO	WO200204493-A2.
PD	17-JAN-2002.
PF	05-JUL-2001; 2001WO-US021241.
PR	05-JUL-2000; 2000US-00610313.
PA	(CHIR ) CHIRON CORP.
PA	(UYST-) UNIV STELLENBOSCH.
PI	Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg EJ;
DR	WPI; 2002-154920/20.
PT	New polynucleotides encoding antigenic HIV Type C polypeptides, useful in
PT	applications including DNA immunization or generation of packaging cell
PT	lines, particularly in gene therapy.
PS	Claim 7; Fig 22; 233pp; English.
XX	
XX	The present invention describes expression cassettes comprising a
CC	polynucleotide sequence encoding a polypeptide comprising immunogenic HIV
CC	type C polypeptides. The expression cassettes comprise any of the HIV
CC	type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or Nef
CC	(I). (I) have immunostimulant activity and can be used in gene therapy.
CC	The HIV type C polynucleotides are useful in applications including DNA
CC	immunisation, generation of packaging cell lines, and production of HIV
CC	Type C proteins. The polynucleotides are particularly useful in gene
CC	therapy and DNA immunisation applications. ABL39942 to ABL40054 and
CC	ABB06204 to ABB06215 represent sequences used in the exemplification of
CC	the present invention
XX	
SQ	Sequence 1494 BP; 325 A; 557 C; 455 G; 157 T; 0 U; 0 Other;

```
Query Match          100.0%; Score 60; DB 6; Length 1494;
Best Local Similarity 100.0%; Pred. No. 1.3e-09;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 GACATCAAGCAGGCGCCCAAGAGAGCCCTTCCGCGACTACGTGAGCCGCTTCTTCAAGACC 60
      |||
      841 GACATCAAGCAGGCGCCCAAGAGAGCCCTTCCGCGACTACGTGAGCCGCTTCTTCAAGACC 900

RESULT 11
ACA03590      ACA03590 standard; DNA; 2742 BP.
AC      ACA03590;
XX      22-MAY-2003 (first entry)
DT
XX      Synthetic DNA encoding immunogenic HIV peptide #73.
DE
XX      Immunogenic HIV polypeptide; human immunodeficiency virus; HIV; vaccine;
KW      gene therapy; packaging cell line; humoral immune response;
KW      cellular immune response; gene delivery vector; DNA immunisation; ds.
XX      Synthetic.
OS      WO2003004657-A1.
XX      16-JAN-2003.
XX      05-JUL-2002; 2002WO-US021421.
PF
XX      05-JUL-2001; 2001US-0303192P.
PR      31-AUG-2001; 2001US-0316860P.
XX      16-JAN-2002; 2002US-0349728P.
PR      16-JAN-2002; 2002US-0349728P.
PR      16-JAN-2002; 2002US-0349793P.
PR      16-JAN-2002; 2002US-0349871P.
XX      (CHIR ) CHIRON CORP.
PA
XX      Zur Megede J, Barnett SW, Lian Y;
PI
XX      WPI; 2003-221602/21.
DR
XX      New synthetic polynucleotides encoding antigenic HIV type B and/or type C
PT      polypeptides, useful as immunogenic compositions or vaccines for
PT      generating humoral or cellular immune responses against HIV in a subject,
PT      especially humans.
XX
XX      Example 1; Fig 78; 262pp; English.
PS
XX      The invention describes a synthetic polynucleotide encoding 2 or more
CC      immunogenic HIV polypeptides, where at least 2 of the polypeptides are
CC      derived from different HIV subtypes. The polynucleotide is useful for
CC      immunisation, generation of packaging cell lines, or production of HIV
CC      polypeptides. The polynucleotide and its encoded proteins are useful as
CC      immunogenic compositions or vaccines for generating humoral or cellular
CC      immune responses against HIV in a subject, or for inducing neutralising
CC      antibodies against HIV. The gene delivery vector comprising the
CC      polynucleotide is also useful for DNA immunisation of, or for generating
CC      an immune response (e.g. a humoral or cellular immune response) in, a
CC      subject such as a mammal, particularly a human. This sequence encodes a
CC      human immunodeficiency virus immunogenic peptide
XX
SQ      Sequence 2742 BP; 578 A; 1020 C; 859 G; 285 T; 0 U; 0 Other;

Query Match          100.0%; Score 60; DB 7; Length 2742;
Best Local Similarity 100.0%; Pred. No. 1.4e-09;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 GACATCAAGCAGGCGCCCAAGAGAGCCCTTCCGCGACTACGTGAGCCGCTTCTTCAAGACC 60
      |||
      2104 GACATCAAGCAGGCGCCCAAGAGAGCCCTTCCGCGACTACGTGAGCCGCTTCTTCAAGACC 2163
DB
```

```
RESULT 12
ACA03524      ACA03524 standard; DNA; 2742 BP.
ID      ACA03524
XX      ACA03524;
AC
XX      22-MAY-2003 (first entry)
DT
XX      Synthetic DNA encoding immunogenic HIV peptide #7.
DE
XX      Immunogenic HIV polypeptide; human immunodeficiency virus; HIV; vaccine;
KW      gene therapy; packaging cell line; humoral immune response;
KW      cellular immune response; gene delivery vector; DNA immunisation; ds.
XX      Synthetic.
OS      WO2003004657-A1.
XX      16-JAN-2003.
XX      05-JUL-2002; 2002WO-US021421.
PF
XX      05-JUL-2001; 2001US-0303192P.
PR      31-AUG-2001; 2001US-0316860P.
XX      16-JAN-2002; 2002US-0349728P.
PR      16-JAN-2002; 2002US-0349793P.
PR      16-JAN-2002; 2002US-0349871P.
XX      (CHIR ) CHIRON CORP.
PA
XX      Zur Megede J, Barnett SW, Lian Y;
PI
XX      WPI; 2003-221602/21.
DR
XX      New synthetic polynucleotides encoding antigenic HIV type B and/or type C
PT      polypeptides, useful as immunogenic compositions or vaccines for
PT      generating humoral or cellular immune responses against HIV in a subject,
PT      especially humans.
XX
XX      Example 1; Fig 12; 262pp; English.
PS
XX      The invention describes a synthetic polynucleotide encoding 2 or more
CC      immunogenic HIV polypeptides, where at least 2 of the polypeptides are
CC      derived from different HIV subtypes. The polynucleotide is useful for
CC      immunisation, generation of packaging cell lines, or production of HIV
CC      polypeptides. The polynucleotide and its encoded proteins are useful as
CC      immunogenic compositions or vaccines for generating humoral or cellular
CC      immune responses against HIV in a subject, or for inducing neutralising
CC      antibodies against HIV. The gene delivery vector comprising the
CC      polynucleotide is also useful for DNA immunisation of, or for generating
CC      an immune response (e.g. a humoral or cellular immune response) in, a
CC      subject such as a mammal, particularly a human. This sequence encodes a
CC      human immunodeficiency virus immunogenic peptide
XX
SQ      Sequence 2742 BP; 578 A; 1020 C; 859 G; 285 T; 0 U; 0 Other;

Query Match          100.0%; Score 60; DB 7; Length 2742;
Best Local Similarity 100.0%; Pred. No. 1.4e-09;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 GACATCAAGCAGGCGCCCAAGAGAGCCCTTCCGCGACTACGTGAGCCGCTTCTTCAAGACC 60
      |||
      850 GACATCAAGCAGGCGCCCAAGAGAGCCCTTCCGCGACTACGTGAGCCGCTTCTTCAAGACC 909
DB
```

```
RESULT 13
ADCl3278      ADCl3278 standard; DNA; 2742 BP.
ID      ADCl3278
XX      ADCl3278;
AC
XX      18-DEC-2003 (first entry)
DT
```

XX DE DNA of HIV construct TatRevNefGag\_C SEQ ID NO 57.  
XX  
XX expression cassette; HIV Gag; Env; Int; Nef; p15RnaseH; Pol; Tat; Prot;  
KW Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; ds.  
XX  
OS Human immunodeficiency virus.  
XX  
PN WO2003004620-A2.  
XX  
XX 16-JAN-2003.  
XX  
XX 05-JUL-2002; 2002WO-US021420.  
XX  
XX 05-JUL-2001; 2001US-0303192P.  
PR 31-AUG-2001; 2001US-0316860P.  
PR 16-JAN-2002; 2002US-0349871P.  
XX  
XX (CHIR ) CHIRON CORP.  
PA (UYST-) UNIV STELLENBOSCH.  
XX  
XX Zur Megede J, Barnett SW, Lian Y, Engelbrecht S, Van Rensburg EJ;  
PI WPI; 2003-221593/21.  
XX  
XX New expression cassette comprising a polynucleotide sequence encoding a  
PT polypeptide including an HIV Gag, Env, Int, Nef, p15RnaseH, Pol, Tat,  
PT Prot, or Rev polypeptide, useful for immunization, or generating  
PT packaging cell lines.  
XX  
XX Disclosure; Fig 54; 301pp; English.  
XX  
XX The invention relates to a novel expression cassette comprising a  
CC polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,  
CC Int, Nef, p15RnaseH, Pol, Tat, Prot, or Rev polypeptide. The novel  
CC expression cassette can be used to treat HIV type C by gene therapy or  
CC used in the development of a vaccine. The gene delivery vector is  
CC administered intramuscularly, intramucosally, intranasally,  
CC subcutaneously, intradermally, transdermally, intravaginally,  
CC intrarectally, orally or intravenously. The expression cassette is useful  
CC for immunisation, generating packaging cell lines and producing HIV  
CC polypeptides. This polynucleotide sequence represents the DNA of an HIV  
CC Type C related sequence of the invention.  
XX  
XX Sequence 2742 BP; 578 A; 1020 C; 859 G; 285 T; 0 U; 0 Other;  
SQ  
Query Match 100.0%; Score 60; DB 9; Length 2742;  
Best Local Similarity 100.0%; Pred. No. 1.4e-09;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GACATCAAGCAGGGCCCCAAGAGAGCCCTTCGCGACTAGTGAGCCGCTTCTCAAGACC 60  
DB 2104 GACATCAAGCAGGGCCCCAAGAGAGCCCTTCGCGACTAGTGAGCCGCTTCTCAAGACC 2163  
RESULT 14  
ADCI3241  
ID ADCI3241 standard; DNA; 2742 BP.  
XX  
XX ADCI3241;  
XX  
XX 18-DEC-2003 (first entry)  
XX  
XX DNA of HIV construct GagTatRevNef\_C SEQ ID NO 20.  
XX  
XX expression cassette; HIV Gag; Env; Int; Nef; p15RnaseH; Pol; Tat; Prot;  
KW Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; ds.  
XX  
XX Human immunodeficiency virus.  
XX  
XX WO2003004620-A2.  
XX  
XX 16-JAN-2003.  
XX  
XX

XX PF 05-JUL-2002; 2002WO-US021420.  
XX  
XX 05-JUL-2001; 2001US-0303192P.  
PR 31-AUG-2001; 2001US-0316860P.  
PR 16-JAN-2002; 2002US-0349871P.  
XX  
XX (CHIR ) CHIRON CORP.  
PA (UYST-) UNIV STELLENBOSCH.  
XX  
XX Zur Megede J, Barnett SW, Lian Y, Engelbrecht S, Van Rensburg EJ;  
PI WPI; 2003-221593/21.  
XX  
XX New expression cassette comprising a polynucleotide sequence encoding a  
PT polypeptide including an HIV Gag, Env, Int, Nef, p15RnaseH, Pol, Tat,  
PT Prot, or Rev polypeptide, useful for immunization, or generating  
PT packaging cell lines.  
XX  
XX Disclosure; Fig 17; 301pp; English.  
XX  
XX The invention relates to a novel expression cassette comprising a  
CC polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,  
CC Int, Nef, p15RnaseH, Pol, Tat, Prot, or Rev polypeptide. The novel  
CC expression cassette can be used to treat HIV type C by gene therapy or  
CC used in the development of a vaccine. The gene delivery vector is  
CC administered intramuscularly, intramucosally, intranasally,  
CC subcutaneously, intradermally, transdermally, intravaginally,  
CC intrarectally, orally or intravenously. The expression cassette is useful  
CC for immunisation, generating packaging cell lines and producing HIV  
CC polypeptides. This polynucleotide sequence represents the DNA of an HIV  
CC Type C related sequence of the invention.  
XX  
XX Sequence 2742 BP; 578 A; 1020 C; 859 G; 285 T; 0 U; 0 Other;  
SQ  
Query Match 100.0%; Score 60; DB 9; Length 2742;  
Best Local Similarity 100.0%; Pred. No. 1.4e-09;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GACATCAAGCAGGGCCCCAAGAGAGCCCTTCGCGACTAGTGAGCCGCTTCTCAAGACC 60  
DB 850 GACATCAAGCAGGGCCCCAAGAGAGCCCTTCGCGACTAGTGAGCCGCTTCTCAAGACC 909  
RESULT 15  
ADCI3239  
ID ADCI3239 standard; DNA; 3162 BP.  
XX  
XX ADCI3239;  
XX  
XX 18-DEC-2003 (first entry)  
XX  
XX DNA of HIV construct GagRtmut\_C SEQ ID NO 18.  
XX  
XX expression cassette; HIV Gag; Env; Int; Nef; p15RnaseH; Pol; Tat; Prot;  
KW Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; ds.  
XX  
XX Human immunodeficiency virus.  
XX  
XX WO2003004620-A2.  
XX  
XX 16-JAN-2003.  
XX  
XX 05-JUL-2002; 2002WO-US021420.  
XX  
XX 05-JUL-2001; 2001US-0303192P.  
PR 31-AUG-2001; 2001US-0316860P.  
PR 16-JAN-2002; 2002US-0349871P.  
XX  
XX (CHIR ) CHIRON CORP.  
PA (UYST-) UNIV STELLENBOSCH.  
XX  
XX Zur Megede J, Barnett SW, Lian Y, Engelbrecht S, Van Rensburg EJ;  
PI

XX  
DR WPI, 2003-221593/21.

XX  
PT New expression cassette comprising a polynucleotide sequence encoding a  
PT polypeptide including an HIV Gag, Env, Int, Nef, p15RnaseH, Pol, Tat,  
PT Prot, or Rev polypeptide, useful for immunization, or generating  
PT packaging cell lines.

XX  
PS Disclosure; Fig 15; 301pp; English.

XX  
CC The invention relates to a novel expression cassette comprising a  
CC polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,  
CC Int, Nef, p15RnaseH, Pol, Tat, Prot, or Rev polypeptide. The novel  
CC expression cassette can be used to treat HIV type C by gene therapy or  
CC used in the development of a vaccine. The gene delivery vector is  
CC administered intramuscularly, intramusosally, intranasally,  
CC subcutaneously, intradermally, transdermally, intravaginally,  
CC intrarectally, orally or intravenously. The expression cassette is useful  
CC for immunisation, generating packaging cell lines and producing HIV  
CC polypeptides. This polynucleotide sequence represents the DNA of an HIV  
CC Type C related sequence of the invention.

XX  
SQ Sequence 3162 BP; 729 A; 1095 C; 975 G; 363 T; 0 U; 0 Other;

Query Match 100.0%; Score 60; DB 9; Length 3162;  
Best Local Similarity 100.0%; Pred. No. 1.4e-09;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATCAAGCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGACCGCTTCTCAAGACC 60  
|||  
850 GACATCAAGCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGACCGCTTCTCAAGACC 909

Search completed: May 28, 2004, 09:49:05  
Job time : 29.8049 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 1, 2004, 11:29:01 ; Search time 141 Seconds  
(without alignments)  
5939.155 Million cell updates/sec

Title: US-09-475-704A-4  
Perfect score: 1509  
Sequence: 1 atggcgcccgccgcagcat.....gcggcccccctgagccagtaa 1509

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 682709 segs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA: \*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq: \*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq: \*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq: \*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq: \*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq: \*  
6: /cgn2\_6/ptodata/2/ina/backfile1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1201	79.6	1515	4	US-09-475-515-4 Sequence 4, Appli
2	1201	79.6	4472	4	US-09-475-515-75 Sequence 75, Appli
3	1201	79.6	4608	4	US-09-475-515-76 Sequence 76, Appli
4	1201	79.6	4689	4	US-09-475-515-74 Sequence 74, Appli
5	1201	79.6	4766	4	US-09-475-515-73 Sequence 73, Appli
6	1199	79.5	2031	4	US-09-475-515-7 Sequence 7, Appli
7	1117.8	74.1	1853	4	US-09-475-515-5 Sequence 5, Appli
8	1117.8	74.1	1865	4	US-09-475-515-78 Sequence 78, Appli
9	1117.8	74.1	1865	4	US-09-475-515-79 Sequence 79, Appli
10	1117.8	74.1	4319	4	US-09-475-515-6 Sequence 6, Appli
11	1049.2	69.5	1268	4	US-09-475-515-9 Sequence 9, Appli
12	929.4	61.6	4307	4	US-09-552-950-2 Sequence 2, Appli
13	925.4	61.3	9772	4	US-09-552-950-5 Sequence 5, Appli
14	736.6	48.8	8366	4	US-09-872-733A-6 Sequence 6, Appli
15	734.2	48.7	1496	4	US-09-184-418C-82 Sequence 82, Appli
16	734.2	48.7	8972	4	US-09-184-418C-9 Sequence 9, Appli
17	723.8	48.0	4338	4	US-09-872-733A-1 Sequence 1, Appli
18	709.6	47.0	1485	4	US-09-184-418C-73 Sequence 73, Appli
19	709.6	47.0	9010	4	US-09-184-418C-8 Sequence 8, Appli
20	708	46.9	1476	4	US-09-184-418C-100 Sequence 100, App
21	708	46.9	8959	4	US-09-184-418C-11 Sequence 11, Appli
22	663.6	44.0	1486	4	US-09-184-418C-38 Sequence 38, Appli
23	663.6	44.0	8992	4	US-09-184-418C-4 Sequence 4, Appli
24	651.8	43.2	8968	4	US-09-184-418C-1 Sequence 1, Appli
25	649.4	43.0	1479	4	US-09-184-418C-12 Sequence 12, Appli
26	642.2	42.6	7399	2	US-08-418-848A-9 Sequence 9, Appli
27	642.2	42.6	9709	2	US-08-188-583-5 Sequence 5, Appli

28	642.2	42.6	9709	3	US-08-388-353-1 Sequence 1, Appli
29	642.2	42.6	9709	3	US-08-488-551B-1 Sequence 1, Appli
30	642.2	42.6	9709	4	US-09-309-572-15 Sequence 15, Appli
31	642.2	42.6	9709	4	US-09-718-096-15 Sequence 15, Appli
32	642.2	42.6	12479	4	US-09-318-138-13 Sequence 13, Appli
33	642.2	42.6	12494	3	US-08-935-312-13 Sequence 13, Appli
34	642.2	42.6	12494	3	US-08-848-760B-33 Sequence 33, Appli
35	642.2	42.6	15581	3	US-08-646-538-35 Sequence 35, Appli
36	642.2	42.6	15581	3	US-09-503-222-35 Sequence 35, Appli
37	633	41.9	1509	4	US-09-475-515-1 Sequence 1, Appli
38	633	41.9	1610	4	US-09-020-144-1 Sequence 1, Appli
39	633	41.9	1845	4	US-09-475-515-2 Sequence 2, Appli
40	633	41.9	4313	4	US-09-475-515-3 Sequence 3, Appli
41	633	41.9	9737	2	US-08-944-449-7 Sequence 7, Appli
42	633	41.9	9737	4	US-09-353-362-7 Sequence 7, Appli
43	632.6	41.9	2025	4	US-09-475-515-8 Sequence 8, Appli
44	631.4	41.8	5362	3	US-08-463-210-5 Sequence 5, Appli
45	631.4	41.8	5362	4	US-08-463-028-5 Sequence 5, Appli

## ALIGNMENTS

RESULT 1  
US-09-475-515-4  
Sequence 4, Application US/09475515A

Patent No. 6602705

GENERAL INFORMATION:

APPLICANT: BARNETT, Susan

APPLICANT: ZUR MEGEDR, Jan

APPLICANT: SRIVASTAVA, Indresh

APPLICANT: LIAN, Ying

APPLICANT: HARTOG, Karin

APPLICANT: LIU, Hong

APPLICANT: GREER, Catherine

APPLICANT: SELBY, Mark

APPLICANT: WALKER, Christopher

TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION

TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES

FILE REFERENCE: 1621.002

CURRENT APPLICATION NUMBER: US/09/475,515A

CURRENT FILING DATE: 1999-12-30

NUMBER OF SEQ ID NOS: 90

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 4

LENGTH: 1515

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURES:

OTHER INFORMATION: Description of Artificial Sequence: synthetic

US-09-475-515-4

Query Match 79.6%; Score 1201; DB 4; Length 1515;

Best Local Similarity 88.5%; Pred. No. 9.3e-170;

Matches 1351; Conservative 0; Mismatches 140; Indels 36; Gaps 3;

QY	1	ATGGGCGCCCGCCGCAATCTGCGCGGAGAGAGCTGAGCAAGTGGAGAATCCGC	60
DB	7	ATGGGCGCCCGCCGCAATCTGAGCGGCGGAGCTGAGCAAGTGGAGAATCCGC	66
QY	61	CTGCGCCCGCGCGGCAAGCACTACATCTGAAGCACTGTGTGGCCAGCCGCGAG	120
DB	67	CTGCGCCCGCGCGGCAAGCACTACATCTGAAGCACTGTGTGGCCAGCCGCGAG	126
QY	121	CTGAGGCGCTTGGCCCTGAACCCCGGCTGTGAGACCCGCGAGGCGTGAAGCATC	180
DB	127	CTGAGGCGCTTGGCCCTGAACCCCGGCTGTGAGACCCGCGAGGCGTGAAGCATC	186
QY	181	ATGAAGCACTGCAAGCCCGCTGCAAGCCGCAAGAGAGCTGCGCAAGCTGTACAAC	240
DB	187	CTGGGCACTGCAAGCCCGCTGCAAGCCGCAAGAGAGCTGCGCAAGCTGTACAAC	246

```
OY 241 ACCGTGGCCACCTCTGTACTGCTGCAACCCGGCATCGAGTCCGCGACCAAGAGGCC 300
DB 247 ACCGTGGCCACCTCTGTACTGCTGCAACCCGGCATCGAGTCCGCGACCAAGAGGCC 306
OY 301 CTGACAGAGATCGAGAGAGACAAAGTCCAGCAGAGAACCCAGCAGGCCCAAGAG 360
DB 307 CTGAGAGAGATCGAGAGAGACAAAGTCCAGAGAGAGGCCCAAGCAGGCCCGCC 366
OY 361 GCCGAGGCA-----AGTGAAGCCAGAACTACCCCATCGTGCAGAACTTG 405
DB 367 GCCGCGGCAACCGGCAACGACGAGCCAGGTGAGCCAGAACTACCCCATCGTGCAGAACTTG 426
OY 406 CAGGCGCAGATGTGTGCAACGAGCCCATAGCCCGCCAGCCCTGAAAGCTGGTGAAGT 465
DB 427 CAGGCGCAGATGTGTGCAACGAGCCCATAGCCCGCCAGCCCTGAAAGCTGGTGAAGT 486
OY 466 ATCGAGAGAGAGGCTTTCAGCCCGGAGTGTATCCCATGTTCAACCGCCCTGAGCGAGGGC 525
DB 487 GTGAGAGAGAGAGGCTTTCAGCCCGGAGTGTATCCCATGTTCAACCGCCCTGAGCGAGGGC 546
OY 526 GCCACCCCGGAGGAGCTTGAACAGAGTGTGAACAACCTGGGGCGGCCACCAAGCCCGCAT 585
DB 547 GCCACCCCGGAGGAGCTTGAACAGAGTGTGAACAACCTGGGGCGGCCACCAAGCCCGCAT 606
OY 586 CAGATGCTGAAGACACCATCAACGAGAGGCGCGGAGTGGGACCGCTGCACCCCGTG 645
DB 607 CAGATGCTGAAGAGACCATCAACGAGAGGCGCGGAGTGGGACCGCTGCACCCCGTG 666
OY 646 CAGGCGGCGCGCGTGGCGCGCGGCAAGTGGCGGACCCCGGCGGAGCGACATCGCCGGC 705
DB 667 CAGGCGGCGCGCGCATCGCGCGCGGCAAGTGGCGGACCCCGGCGGAGCGACATCGCCGGC 726
OY 706 GCCACGAGCAACCTGACAGAGCAGATCGCTGATGACAGCAACCCCGCTGCGCGTG 765
DB 727 AGCAGCAGCAACCTGACAGAGCAGATCGCTGATGACAGCAACCCCGCTGCGCGTG 786
OY 766 GCGGACATCTACAAGCGGTGATCATCTGGGCTGAAACAAGATGTGCGATGTACAGC 825
DB 787 GCGGAGATCTACAAGCGGTGATCATCTGGGCTGAAACAAGATGTGCGATGTACAGC 846
OY 826 CCGGTGAGCATCTGACATCGCGCAGAGCCCAAGAGCCCTTCGCGCACTACGTGAC 885
DB 847 CCGCAGCAGCATCTGACATCGCGCAGAGCCCAAGAGCCCTTCGCGCACTACGTGAC 906
OY 886 CGCTTCTTCAAGACCTGCGCGCGGAGAGGCAACCCAGAGCGTGAAGAATGTGATGAC 945
DB 907 CGCTTCTTCAAGACCTGCGCGCGTGAAGAGGCAACCCAGAGCGTGAAGAATGTGATGAC 966
OY 946 GAGACCTGTGTGTGCAAGCGCAACCCGCACTGCAAGACCATCTGCGCGCTCTCGC 1005
DB 967 GAGACCTGTGTGTGCAAGCGCAACCCGCACTGCAAGACCATCTGAGGCTCTCGC 1026
OY 1006 CCGGCGGCAACCTGAGAGAGATGATGACCGCTGCAAGGCGTGGCGCGCCCGGCAC 1065
DB 1027 CCGGCGGCAACCTGAGAGAGATGATGACCGCTGCAAGGCGTGGCGCGCCCGGCAC 1086
OY 1066 AAGGCGCGGTGTGTGCGGAGCGAGTGAAGCAAG---CCAAGCGTGAACATCATGATG 1122
DB 1087 AAGGCGCGGTGTGTGCGGAGCGAGTGAAGCAAGCCCGGCAACATCATGATG 1146
OY 1123 CAGAGAGCAACTTCAAGGGCCCCCGGCAAGCTCAAGTCTTCAACTGCGGCAAGAG 1182
DB 1147 CAGGCGGCAACTTCCGCAACGAGAGAACCGTCAAGTCTTCAACTGCGGCAAGAG 1206
OY 1183 GGCACATCGCCAAAGACTGCGCGCGCCCGCAAGAGGCTGCTGAAGTGCAGAG 1242
DB 1207 GGCACACCGCCAGAGACTGCGCGCGCCCGCAAGAGGCTGCTGCGCTGCGCGC 1266
OY 1243 GAGGCGCACAGATGAAGACTGCAACGAGCGCAGGCACTTCTGGGCAAGATCTGG 1302
DB 1267 GAGGCGCACAGATGAAGACTGCAACGAGCGCAGGCACTTCTGGGCAAGATCTGG 1326
OY 1303 CCGAGCAACAAGGCGCGCCCGCAACTTCTGCAAGACCGCAAGCGCGCGCGCC 1362
```

```
DB 1327 CCCAGCTACAAGGCGCGCCCGGCAACTTCTGCAAGCGC-----C 1368
OY 1363 ACCGTGCCACCGCCCGCGGAGAGCTTCGCTTGAAGAGACCAACCCCGCC 1422
DB 1369 CCGAGCCCAACCGCCCGCGGAGAGAGCTTCGCTTGAAGAGAGACCAACCC 1428
OY 1423 AAGCAGAGCCCAAGAGAGCGAGCCCTTACCGGAGCCCTGACCGCCCTGCGAGCTG 1482
DB 1429 AGCCAGAGCAGAGAGCCATGACAGAGAGCTGTACCTCTGACAGCTGCGAGCTG 1488
OY 1483 TTCGCAAGCGCGCCCTGAGCCAGTAA 1509
DB 1489 TTCGCAAGCAGCCCGAGCAGCACTAA 1515
```

RESULT 2  
US-09-475-515-75  
Sequence 75, Application US/09475515A  
Patent No. 6602705

GENERAL INFORMATION:  
APPLICANT: BARRETT, Susan  
APPLICANT: ZUR MEGEDE, Jan  
APPLICANT: SRIVASTAVA, Indresh  
APPLICANT: LIAN, Ying  
APPLICANT: HARTOG, Karin  
APPLICANT: LIU, Hong  
APPLICANT: GREER, Catherine  
APPLICANT: SELBY, Mark  
APPLICANT: WALKER, Christopher  
TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION  
TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES  
FILE REFERENCE: 1621.002  
CURRENT APPLICATION NUMBER: US/09/475,515A  
CURRENT FILING DATE: 1999-12-30  
NUMBER OF SEQ ID NOS: 90  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 75  
LENGTH: 4472  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURES:  
OTHER INFORMATION: Description of Artificial Sequence:  
US-09-475-515-75

Query Match 79.6%; Score 1201; DB 4; Length 4472;  
Best Local Similarity 88.5%; Pred. No. 9e-170;  
Matches 1351; Conservative 0; Mismatches 140; Indels 36; Gaps 3;

```
OY 1 ATGGGCGCGCGCCAGATCTGCGCGGAGAGAGCTGCAAGTGGAGAAAGATCCGC 60
DB 2939 ATGGGCGCGCGCCAGCGTGTGAGCGCGGAGAGCTGCAAGTGGAGAAAGATCCGC 2998
OY 61 CTGCGCCCGCGCGCAAGAGCACTTACATGCTGAAGCACCTGTGTGGCCAGCCGAG 120
DB 2999 CTGCGCCCGCGCGCAAGAGAGTACAGCTGAAGCATGTGTGGCCAGCCGAG 3058
OY 121 CTGAGGCTTGCCTTGAACCCCGCTGCTGAGACCGCGAGGGCTGCAAGCATC 180
DB 3059 CTGAGGCTTGCCTTGAACCCCGCTGCTGAGACCGCGAGGGCTGCGCAGATC 3118
OY 181 ATGAAGAGCTGACGCGCCCTGCAAGACCGGCAACGAGAGCTGCGAGCTGTACAAC 240
DB 3119 CTGGGCAAGCTGACGCGCCCTGCAAGACCGGCAACGAGAGCTGCGAGCTGTACAAC 3178
OY 241 ACCGTGCCACCTGTATCTGCTGACGCGCGCATGAGGTCCGCAACCAAGAGGCC 300
DB 3179 ACCGTGCCACCTGTATCTGCTGACGCGCGCATGAGGTCCGCAACCAAGAGGCC 3238
OY 301 CTGACAGAGATCGAGAGAGCAGAACTGCAAGAGAGAGAGAGAGAGAGAGAGAG 360
DB 3239 CTGAGAGAGATCGAGAGAGCAGAACTGCAAGAGAGAGAGAGAGAGAGAGAGAG 3298
```

```
QY 361 GCCGACGGCA-----AGGTGAGCCAGAACTACCCCATCTGTGCAGAACTG 405
DB 3299 GCCCGCCGACCCGGCAACAGCAGCCAGGTGAGCCAGAACTACCCCATCTGTGCAGAACTG 3358
QY 406 CAGGSCCAGATGTGCAACAGGSCCATCAGCCCCCGCACTGAACGCTGGGTGAAGTG 465
DB 3359 CAGGSCCAGATGTGCAACAGGSCCATCAGCCCCCGCACTGAACGCTGGGTGAAGTG 3418
QY 466 ATCGAGGAGAAAGGCTTCAGCCCCGAGGTGATCCCACTGTTCAACCGCTGAGCGAGGC 525
DB 3419 GTGAGGAGAGAAAGGCTTCAGCCCCGAGGTGATCCCACTGTTCAACCGCTGAGCGAGGC 3478
QY 526 GCCACCCCCCAGGACCTGAACAAGATGTGAACAACCGTGGGCGGCAACAGCCGCGCATG 585
DB 3479 GCCACCCCCCAGGACCTGAACAAGATGTGAACAACCGTGGGCGGCAACAGCCGCGCATG 3538
QY 586 CAGATGCTGAAGGACCATCAACGAGGAGCGCCGCGAGTGGGACCGCTGCAACCCCGTG 645
DB 3539 CAGATGCTGAAGGACCATCAACGAGGAGCGCCGCGAGTGGGACCGCTGCAACCCCGTG 3598
QY 646 CAGCGCGCGCGCGCGCGCGCGCGCGAGATGCGGACCGCGCGCGAGCATCGCGCGC 705
DB 3599 CAGCGCGCGCGCGCGCGCGCGCGCGAGATGCGGACCGCGCGCGAGCATCGCGCGC 3658
QY 706 GCCACCAAGCACCCTGCAAGAGAGAGATGCTGATGAACAGCAACCCCGCTGCGCGC 765
DB 3659 ACCACCAAGCACCCTGCAAGAGAGAGATGCTGATGAACAGCAACCCCGCTGCGCGC 3718
QY 766 GGGCAGATCTACAAGCGGTGATCATCTGGGCTGAACAAGATCGTGGGATGTACAGC 825
DB 3719 GGGCAGATCTACAAGCGGTGATCATCTGGGCTGAACAAGATCGTGGGATGTACAGC 3778
QY 826 CCGGTGAGCATCTCTGACATCCGCGCAAGGCGCCCAAGAGCCCTTCCCGCACTACGTGAC 885
DB 3779 CCGGTGAGCATCTCTGACATCCGCGCAAGGCGCCCAAGAGCCCTTCCCGCACTACGTGAC 3838
QY 886 CGCTTCTTCAAGACCTCTGCGCGCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 945
DB 3839 CGCTTCTTCAAGACCTCTGCGCGCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3898
QY 946 GAGACCTCTGCTGTGTCAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1005
DB 3899 GAGACCTCTGCTGTGTCAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3958
QY 1006 CCGCGCGCGCGCGCGCGCGCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1065
DB 3959 CCGCGCGCGCGCGCGCGCGCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4018
QY 1066 AAGGCGCGCGCGCGCGCGCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1122
DB 4019 AAGGCGCGCGCGCGCGCGCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4078
QY 1123 CAGAAAGGACACTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1182
DB 4079 CAGAAAGGACACTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4138
QY 1183 GGGCAGCATCGCAAGAACTGCGCGCGCGCGCGCAAGAAAGGAGGAGGAGGAGGAGGAGGAG 1242
DB 4139 GGGCAGCATCGCAAGAACTGCGCGCGCGCGCGCAAGAAAGGAGGAGGAGGAGGAGGAGGAG 4198
QY 1243 GAGGCGCAACAGATGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1302
DB 4199 GAGGCGCAACAGATGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4258
QY 1303 CCGAGCCCAAGAGGCGCGCGCGCGCAACTTCTGCAAGAACTGCAAGCGAGCGCGCGCGCC 1362
DB 4259 CCGAGCCCAAGAGGCGCGCGCGCGCAACTTCTGCAAGAACTGCAAGCGAGCGCGCGCGCC 4300
QY 1363 ACCGTGCGCGCGCGCGCGCGCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1422
DB 4301 CCGAGCGCGCGCGCGCGCGCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4360
```

```
QY 1423 AAGCAGAGCCCAAGAGCCGAGCCCTACCGGAGCCCTGACCCGCTGCGCAGCCTG 1482
DB 4361 AGCCAGAGCAGAGAGCCCATCGACAGAGAGCTGTACCCCTGACAGCCTGCGCAGCCTG 4420
QY 1483 TTCGAGCGCGCGCGCGCGCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1509
DB 4421 TTCGAGAGAGAGCCCGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4447
```

## RESULT 3

```
US-09-475-515-76
; Sequence 76, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: LIU, Hong
; APPLICANT: GRBER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475, 515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 76
; LENGTH: 4608
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: gp160.modSF162.delV2.gag.modSF2
US-09-475-515-76
```

```
Query Match 79.6%; Score 1201; DB 4; Length 4608;
Best Local Similarity 88.5%; Pred. No. 9e-170;
Matches 1351; Conservative 0; Mismatches 140; Indels 36; Gaps 3;

QY 1 ATGGCGCGCGCGCGCGCGCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 60
DB 3075 ATGGCGCGCGCGCGCGCGCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3134
QY 61 CTGCGCGCGCGCGCGCGCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
DB 3135 CTGCGCGCGCGCGCGCGCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3194
QY 121 CTGAGGAGGCTTGGCCCTGAACCCCGGCTGCTGAGAGCCCGGAGGAGGAGGAGGAGGAG 180
DB 3195 CTGAGGAGGCTTGGCCCTGAACCCCGGCTGCTGAGAGCCCGGAGGAGGAGGAGGAGGAG 3254
QY 181 ATGAAGCAGCTGCAAGCGCGCGCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
DB 3255 CTGAGGAGGCTGCAAGCGCGCGCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3314
QY 241 ACCGTGCGCGCGCGCGCGCGCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
DB 3315 ACCGTGCGCGCGCGCGCGCGCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3374
QY 301 CTGAGCAGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
DB 3375 CTGAGCAGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3434
QY 361 GCCGAGCGCA-----AGGTGAGCCAGAACTACCCCATCTGTGCAAGAACTG 405
DB 3435 GCCGAGCGCA-----AGGTGAGCCAGAACTACCCCATCTGTGCAAGAACTG 3494
QY 406 CAGGCGCAGATGTGCAACAGGSCCATCAGCCCCCGCACTGAAACGCTGGGTGAAGTG 465
```

Db	3495	CAGGGCCAGATGCTGACCAAGGCATCAGCCCCCGCACCCCTGAACGCTGGGTGAAGGTG	3554
Oy	466	ATCGAGAGAAAGGCTTTCAGCCCCGAGGTGATCCCATGTTCACCGCCCTGAGCGAAGGC	525
Db	3555	GTGAGAGAAAGGCTTTCAGCCCCGAGGTGATCCCATGTTCAGCGCCCTGAGCGAAGGC	3614
Oy	526	GCCACCCCCCAGGACCTGAACAAGATGTTGAACAAGTGGGCGGCCACAGGCGGCATG	585
Db	3615	GCCACCCCCCAGGACCTGAACAAGATGTTGAACAAGTGGGCGGCCACAGGCGGCATG	3674
Oy	586	CAGATGCTGAAGGACACCATCAACGAGAGGCGCCGAGTGGGACCGCTGCACCCCGTG	645
Db	3675	CAGATGCTGAAGGACACCATCAACGAGAGGCGCCGAGTGGGACCGCTGCACCCCGTG	3734
Oy	646	CAGGCGGCCCCGTGGCCCCCGGCGCAGATGCGCGACCCCGCGCAGCGACATCGCCGC	705
Db	3735	CAGGCGGCCCCCATCGCCCCCGGCGCAGATGCGCGACCCCGCGCAGCGACATCGCCGC	3794
Oy	706	GCCACCAAGACCTTGCAAGAGCAGATCGCTTGAAGAACAGCAACCCCGCTGCCCCGTG	765
Db	3795	ACCACCAAGACCTTGCAAGAGCAGATCGCTTGAAGAACCAACCCCGCATCCCGTG	3854
Oy	766	GGCGCACTCTACAAAGCGGTGATCATCTGGGCTGAACAAAGATCGTGGGATGTACAGC	825
Db	3855	GGCGGATCTACAAAGCGGTGATCATCTGGGCTGAACAAAGATCGTGGGATGTACAGC	3914
Oy	826	CCCGTGAACATCTTGACATTCGCCAGGGGCCCAAGAGCCCTTCCGCGACTACGTGAC	885
Db	3915	CCCAACCAACATCTTGACATTCGCCAGGGGCCCAAGAGCCCTTCCGCGACTACGTGAC	3974
Oy	886	CGCTTCTTCAAGACCCCTGCGGCCGAGCAGGCCACCCAGAGCGTGAAGAACTGGATGAC	945
Db	3975	CGCTTCTTCAAGACCCCTGCGCGCTGAGCAGGCCACCCAGAGCGTGAAGAACTGGATGAC	4034
Oy	946	GAGACCTGCTGGTGCAGAACGCCCAACCCCGACTGCAAGAACCATCTGCGCGCTCTCGC	1005
Db	4035	GAGACCTGCTGGTGCAGAACGCCCAACCCCGACTGCAAGAACCATCTGCGCGCTCTCGC	4094
Oy	1006	CCCGCGCCCAACCTTGAGAGATGATGACCGGCTGCCAGGGCGTGGGCGGCCCGGCCAC	1065
Db	4095	CCCGCGCCCAACCTTGAGAGATGATGACCGGCTGCCAGGGCGTGGGCGGCCCGGCCAC	4154
Oy	1066	AAGGCCCGCTGCTGGCCGAGGCGATGAGCCAGG---CCAAACGCGTGAACATCATGATG	1122
Db	4155	AAGGCCCGCTGCTGGCCGAGGCGATGAGCCAGGTCAGGAACCCCGCGAACCATCATGATG	4214
Oy	1123	CAGAAAGCAACTTCMAAGGGCCCCCGGCGCAACGTCMAAGTCTTCAACTGCGGCAAGAG	1182
Db	4215	CAGCGCGCAACTTCGCGCAACAGCGGAAGACCGTCMAAGTCTTCAACTGCGGCAAGAG	4274
Oy	1183	GGCCACATCGCCCAAGAACTGCGCGGCCCGCCCGCAAGAGGGCTGCTGGAAGTGGCGCAAG	1242
Db	4275	GGCCACATCGCCCAAGAACTGCGCGGCCCGCCCGCAAGAGGGCTGCTGCGCGCGC	4334
Oy	1243	GAGGGCCACAGATGAAGGACTGCAACGAGCGCCAGGCCCAACTTCTTGGGCAAGATCTGG	1302
Db	4335	GAGGGCCACAGATGAAGGACTGCAACGAGCGCCAGGCCCAACTTCTTGGGCAAGATCTGG	4394
Oy	1303	CCCAAGCCAAAGGGCGGCCCGCGCAACTTCTTGCAAGAACCGCAGCGAGCCCGCGCCCC	1362
Db	4395	CCCAAGCTAACAGGGCGGCCCGCGCAACTTCTTGCAAGAGCCG-----C	4436
Oy	1363	ACCGTGCCCAACGCCCCCGCGCGAGAGCTTCCGCTTGAGAGAGACACCCCGCCCCC	1422
Db	4437	CCCGAGCCCAACGCCCCCGCGCGAGAGAGCTTCCGCTTGCGGAGAGAGAACACACCCCC	4496
Oy	1423	AAGCAGAGCCCAAGGACCGCGAGCCCTTACCGCGAGCCCTTGAACCGCCCTGCGAGCCTG	1482
Db	4497	AGCCAGAGACAGAGAGCCCATGCAAAAGAGCTGTAACCCCTGACCAAGCCTTGCAGCCTG	4556
Oy	1483	TTGCGCAGCGGCCCCCTGAGCCAGTAA	1509
Db	4557	TTGCGCAACGACCCCGAGCAGCCAGTAA	4583

```

RESULT 4
; US-09-475-515-74
; Sequence 74, Application US/09475515A
; Patent No. 6602705
;
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 74
; LENGTH: 4689
;
; TYPE: DNA
; ORGANISM: Artificial Sequence
;
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: gp160.modsF162.gag.modsF2
; US-09-475-515-74

```

Query Match	79.6%	Score 1201	DB 4	Length 4689
Best Local Similarity	88.5%	Pred. No. 9e-170		
Matches 1351	Conservative	0	Mismatches 140	Indels 36
				Gaps 3
Qy	1	ATGGCGCCCGCGCGCAGCATCCTGCGCGCGGAGAGCTGTGACAAAGTGGGAGAAATCCGC		60
Db	3156	ATGGCGCCCGCGCGCAGCGCTGTGAGCGGCGGAGCTGTGACAAAGTGGGAGAAATCCGC		3215
Qy	61	CTGCGCCCGCGCGCGCAGAGCACTACATGCTGAGCACCTGTGTGGGCCAGCCGCGAG		120
Db	3216	CTGCGCCCGCGCGCGCAGAGCAAGTACAACTGAGCACATCTGTGTGGGCCAGCCGCGAG		3275
Qy	121	CTGAGAGGCTTTCGCTCTGAAACCCCGGCTGTCTGTGAGACCGCCGAGGGCTGTGACAGATC		180
Db	3276	CTGAGAGGCTTTCGCTCTGAAACCCCGGCTGTCTGTGAGACCGAGGGCTGTGACAGATC		3335
Qy	181	ATGAAGCAGCTGCAGCCCGGCTGTGACACCGGACCCGAGAGCTGTGCGCAGCTGTACAAC		240
Db	3336	CTGGCGCAGCTGCAGCCCGGCTGTGACACCGGACCGGAGAGCTGTGCGCAGCTGTACAAC		3395
Qy	241	ACCGTGGCCACCTGTACTGCGGTGCACGCGGCGCATGAGGTTCGCGACACCAAGAGGCC		300
Db	3396	ACCGTGGCCACCTGTACTGCGGTGCACAGCGGCGCATGAGCTCAAGGACACCAAGAGGCC		3455
Qy	301	CTGACAAAGATCGAGAGAGGACAAACAACTTCCAGCAGAGAACCCAGCAGGCCCAAGAG		360
Db	3456	CTGAGAAAGATCGAGAGAGGACAAACAACTTCCAGAGAGAGGCCAGCAGGCCCGCGCC		3515
Qy	361	GCCGACGGCA-----AGGTGAGCCAGAACTACCCCATCGTGCAGAACTTG		405
Db	3516	GCCGACGGCAACCGGCAACAGACGACGAGGTGAGCCAGAACTACCCCATCGTGCAGAACTTG		3575
Qy	406	CAGGGCCAGATGGTGCACCAAGGCCCATYACGCTCCCGCACCCCTGAAAGCGCTGGGTGAAGTG		465
Db	3576	CAGGGCCAGATGGTGCACCAAGGCCCATYACGCTCCCGCACCCCTGAAAGCGCTGGGTGAAGTG		3635
Qy	466	ATCGAGAGAGAGGCTTTCAGCTCCCGAGGTGATCCCAATGTTTCAACCGCCTTGAAGCGAGGC		525
Db	3636	GTGAGAGAGAGGCTTTCAGCTCCCGAGGTGATCCCAATGTTTCAAGCGCCTTGAAGCGAGGC		3695
Qy	526	GCCACCCCGCAGGACCTGAAACGATGTTGAACAACGTTGGCGGCGCACCAAGCCGCCATG		585









```
QY 766 GCGCATCTACAAGCGGTGATCATCTGGGCTTGAACAAGATCGTGGATGTACAGC 825
DB 787 GCGGAGATCTACAAGCGGTGATCATCTGGGCTTGAACAAGATCGTGGATGTACAGC 846
QY 826 CCGGTGAGCATCTGAGCATCCGCAAGGCCCCAAGAGGCCCTTCCGCACTACGTGAC 885
DB 847 CCACCAAGCATCTGAGCATCCGCAAGGCCCCAAGAGGCCCTTCCGCACTACGTGAC 906
QY 886 CGCTCTTCAAGAACCTGCGCGCGGAGAGGCCACCCAGAGCTGAAGAACTGATGACC 945
DB 907 CGCTCTTCAAGAACCTGCGCGCGGAGAGGCCACCCAGAGCTGAAGAACTGATGACC 966
QY 946 GAGACCTGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1005
DB 967 GAGACCTGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1026
QY 1006 CCGGCGCCACCTGAGAGAGATGATGACCGCTGCAAGGCGTGGAGCGCCCGGAC 1065
DB 1027 CCGGCGCCACCTGAGAGAGATGATGACCGCTGCAAGGCGTGGAGCGCCCGGAC 1086
QY 1066 AAGGCGCGGTGCTGGCGAGGCGATGAGCAGG---CGAAGAGTGAACATCATGATG 1122
DB 1087 AAGGCGCGGTGCTGGCGAGGCGATGAGCAGG---CGAAGAGTGAACATCATGATG 1146
QY 1123 CAGAAGAGCACTTCAAGGCGCGCGCGCAACCTCAAGTCTTCACTGCGCAAGAG 1182
DB 1147 CAGCGCGCACTTCCGCAACAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1206
QY 1183 GCGCACTGCGCAAGAACTGCGCGCGCGCGCGCAAGAGGCTGCTGAAGTGGGCAAG 1242
DB 1207 GCGCACTGCGCAAGAACTGCGCGCGCGCGCGCAAGAGGCTGCTGAAGTGGGCGCG 1266
QY 1243 GAGGCGCACTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1302
DB 1267 GAGGCGCACTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1326
QY 1303 CCGAGCCCAAGAGGCGCGCGCGCGCACTTCTCTGCAAGAACCGAGCGCGCGCGCGCG 1362
DB 1327 CCGAGCCCAAGAGGCGCGCGCGCGCACTTCTCTGCAAGAGCGG-----C 1368
QY 1363 ACCGTGCGCAACCGCGCGCGCGCGCGAGAGTTCGCTTCAAGAGAGAGAGAGAGAGAG 1422
DB 1369 CCGAGCGCAACCGCGCGCGCGCGCGAGAGAGTTCGCTTCAAGAGAGAGAGAGAGAGAG 1428
QY 1423 AAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1482
DB 1429 AAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1488
QY 1483 TTGGGAGAGCGCGCGCGCGCGAGT 1507
DB 1489 TTGGGAGAGCGAGCGCGCGCGAGT 1513

RESULT 7
US-09-475-515-5
; Sequence 5, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
```

```
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1853
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; US-09-475-515-5

Query Match          74.1%; Score 1117.8; DB 4; Length 1853;
Best Local Similarity 85.1%; Pred. No. 1.8e-157;
Matches 1299; Conservative 0; Mismatches 192; Indels 36; Gaps 3:

QY 1 ATGGCGCGCGCGCGAGCAATCTGCGCGCGAGAGCTGCAAGTGGAGAGATCCGC 60
DB 7 ATGGCGCGCGCGCGAGCGCTGAGCGCGCGAGCTGCAAGTGGAGAGATCCGC 66
QY 61 CTGCGCGCGCGCGAGAGCACTACATCTGAAGCACTGTGTGGAGCGCGAG 120
DB 67 CTGCGCGCGCGCGAGAGAGTACAGCTGAAGCACTGTGTGGAGCGCGAG 126
QY 121 CTGAGAGGCTTGGCTTGAACCCCGGCTGCTGAGAGACCGCGAGGCTGCAAGCATC 180
DB 127 CTGAGAGGCTTGGCTTGAACCCCGGCTGCTGAGAGACCGCGAGGCTGCGCAGATC 186
QY 181 ATGAAGAGCTGAGCGCGCGCTGAGAGACCGAGAGCTGCGAGCTGTACAAC 240
DB 187 CTGCGCGAGCTGAGCGCGAGCTGAGAGACCGAGAGCTGCGAGCTGTACAAC 246
QY 241 ACCGTGCGCACTGTACTGCTGAGAGAGCGCGCATGAGGTCGCGACAGAGAGGCC 300
DB 247 ACCGTGCGCACTGTACTGCTGAGAGAGCGCGCATGAGGTCGCGACAGAGAGGCC 306
QY 301 CTGAGCAAGATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
DB 307 CTGAGAGAGATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 366
QY 361 GCGGAGCGCA-----AGTGAAGCAGAACTACCCCATGCTGAGAACTG 405
DB 367 GCGGAGCGCAACCGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 426
QY 406 CAGGCGCAGATGTGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 465
DB 427 CAGGCGCAGATGTGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 486
QY 466 ATCGAGAGAGAGGCTTCAAGCGCGGAGAGTATCCCATGTTCAACCGCGCTGAGCGAGG 525
DB 487 GTGAGAGAGAGAGGCTTCAAGCGCGGAGAGTATCCCATGTTCAAGCGCGCTGAGCGAGG 546
QY 526 GCCACCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 585
DB 547 GCCACCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 606
QY 586 CAGATGCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 645
DB 607 CAGATGCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 666
QY 646 CAGGCGCGCGCGCGTGGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 705
DB 667 CAGGCGCGCGCGCGTGGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 726
QY 706 GCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 765
DB 727 ACCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 786
QY 766 GCGGAGATCTACAAGCGGTGATCATCTGGGCTTGAACAAGATCGTGGATGTACAGC 825
DB 787 GCGGAGATCTACAAGCGGTGATCATCTGGGCTTGAACAAGATCGTGGATGTACAGC 846
QY 826 CCGGTGAGCATCTGAGCATCCGCAAGGCCCCAAGAGGCCCTTCCGCACTACGTGAC 885
```

```
Db      847 CCCACCAGCATCTGACATCCGCCAGGGCCCCCAAGAGCCCTTCGCGACTACGTGAC 906
QY      886 CGCTTCTCAAGACCTTGGCGGCCGAGCAGGCCCAACCAAGACGTGAAGAACTGGATGACC 945
Db      907 CGCTTCTCAAGACCTTGGCGGCCGAGCAGGCCCAACCAAGACGTGAAGAACTGGATGACC 966
QY      946 GAGACCTGCTGTGACAGAAAGCCCAACCCCGACTGCAAGACCATCTGCGCGCTCTGGGC 1005
Db      967 GAGACCTGCTGTGACAGAAAGCCCAACCCCGACTGCAAGACCATCTGAGGCTCTGGGC 1026
QY      1006 CCCGCGCCCACTTGGAGAGATGATGACCCGCTGCGAGGCGTGGCGGCCCGGCCAC 1065
Db      1027 CCCGCGCCCACTTGGAGAGATGATGACCCGCTGCGAGGCGTGGCGGCCCGGCCAC 1086
QY      1066 AAGCCCGCGTGTGCGCGAGGCGATGAGCCAGG---CCAAGCGTGAACATCATGATG 1122
Db      1087 AAGCCCGCGTGTGCGCGAGGCGATGAGCCAGGACCGCGGCAATCATGATG 1146
QY      1123 CAGAAGACCACTTCAAGGCGCCCGCGCAACGTCAAGTCTTCAACTGCGGCAAGGAG 1182
Db      1147 CAGCGCGCACTTCCGCAACCAAGCGAAGACCGTCAAGTCTTCAACTGCGGCAAGGAG 1206
QY      1183 GGGCAATGCGCAAGAACTGCGCGCGCGCGCGCGCAAGAGGCGTGTGAAGTGGCGCAAG 1242
Db      1207 GGGCAATGCGCGCAAGAACTGCGCGCGCGCGCGCGCGCAAGAGGCGTGTGGCGCGCGC 1266
QY      1243 GAGGCGCAACGATGAAGACTGACCGGCGCGCAAGGCGCAACTTCTGCGCAAGATCTGG 1302
Db      1267 GAAGACACCAATGAAGATGCACTGAGAGACAGGCTAATTTTAGGGAAGATCTGG 1326
QY      1303 CCCAGCCCAAGGGCGCGCGCGCGCAACTTCTGCAAGACCGCAAGCGCGCGCGCGCGCC 1362
Db      1327 CCTTCTCAAGAGGAAGGCGCGGGAATTTTCTCAGAGCAG-----A 1368
QY      1363 AGCGTGCCCAACCGCGCGCGCGCGCGAGAGCTTCCGCTTCAAGAGACCAACCGCGCGCC 1422
Db      1369 CCAGAGCCCAAGCGCGCGCGCGCAAGAGAGAGCTTCAAGTTGGGAGAGAGAAACAATCCC 1428
QY      1423 AAGCAGAGCCCAAGAGACCGCGAGCCCTACCGCGAGCCCTGACCGCGCGCGAGCCTG 1482
Db      1429 TCTCAGAGCAGAGCGCGATAGCAAGGAATGTATCTTAACTTCCCTCAGATCACTC 1488
QY      1483 TTGGCAGCGCGCGCGCGCGAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCA 1509
Db      1489 TTTGGCAAGACCCCTCGTCAAGATAA 1515

RESULT 8
US-09-475-515-78
; Sequence 78, Application US/09475515A
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGBDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GRBBR, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 78
; LENGTH: 1865
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
```

```
; OTHER INFORMATION: Description of Artificial Sequence: GPI
US-09-475-515-78
Query Match      74.1%; Score 1117.8; DB 4; Length 1865;
Best Local Similarity 85.1%; Pred. No. 1.8e-157;
Matches 129; Conservative 0; Mismatches 192; Indels 36; Gaps 3;

QY      1 ATGGCGCGCGCGCGCAGCATCTGCGCGCGAGAGAGCTGGAACAAGTGAGAGAAATCCGC 60
Db      13 ATGGCGCGCGCGCGCAGCATCTGAGCGCGCGAGCTGGAACAAGTGAGAGAAATCCGC 72
QY      61 CTGCGCGCGCGCGCGCAAGAGCACTACATGCTGAAGCACTGTGTGGCCAGCGCGAG 120
Db      73 CTGCGCGCGCGCGCGCAAGAGCACTACATGCTGAAGCACTGTGTGGCCAGCGCGAG 132
QY      121 CTGAGAGGCTTCCCTGTAACCCCGCGCTGCTGAGACCGCGAGGCTGCAAGCATC 180
Db      133 CTGAGAGGCTTCCCTGTAACCCCGCGCTGCTGAGACCGCGAGGCTGCGCATATC 192
QY      181 ATGAAGCAGCTGACCGCGCGCTGCAAGCGCGCACCGAGAGCTGCGACCTGTACAC 240
Db      193 CTGCGCGCAGCTGACCGCGCGCTGCAAGCGCGCACCGAGAGCTGCGACCTGTACAC 252
QY      241 ACCGTGCGCACTCTGTACTGCTGTGCAAGCGCGCATGAGGTCCGCAACCAAGAGGCC 300
Db      253 ACCGTGCGCACTCTGTACTGCTGTGCAAGCGCGCATGAGGTCAAGCAACCAAGAGGCC 312
QY      301 CTGGAACAAGATCGAGAGAGAGAGCAAGTCCAGCAGAGAACCCAGCAGGCCAAGAG 360
Db      313 CTGGAAGAAGATCGAGAGAGAGAGCAAGTCCAGAGAGAGGCCAAGCAGGCCCGCGCC 372
QY      361 GCGGACGGCA-----AGGTAGCCAGAGACTACCCCATGTGTGAAGAACTCTG 405
Db      373 GCGGCGCGCGCGCGCAACAGCAGCAGGTGAGCAGAACTACCCCATGTGTGAAGAACTCTG 432
QY      406 CAGGCGCAAGTGTGCAACGAGCCATCAGCCCGCGCACCTGAAAGCCTGGGTGAAGGTG 465
Db      433 CAGGCGCAAGTGTGCAACGAGCCATCAGCCCGCGCACCTGAAAGCCTGGGTGAAGGTG 492
QY      466 ATCGAGAGAGAGGCTTCAAGCCCGAGGTGATCCCATGTTCACCGCGCTGAGCAGAGGC 525
Db      493 CTGAGAGAGAGAGGCTTCAAGCCCGAGGTGATCCCATGTTCACCGCGCTGAGCAGAGGC 552
QY      526 GCCAAGCGCGCGCGCGCTGAACAGATGTTGAACACCGTGGCGCGCACAGGCGCGCATG 585
Db      553 GCCAAGCGCGCGCGCGCTGAACAGATGTTGAACACCGTGGCGCGCACAGGCGCGCATG 612
QY      586 CAGATGCTGAAGAGCAACCATCAACAGAGAGCGCGCGAGTGGACCGCGCTGACCCCGTG 645
Db      613 CAGATGCTGAAGAGAGCAACCATCAACAGAGAGCGCGCGAGTGGACCGCGCTGACCCCGTG 672
QY      646 CAGGCGCGCGCGCGTGGCGCGCGCGCGCGAGTGGCGGACCGCGCGCGCGCATGCGCGGC 705
Db      673 CAGGCGCGCGCGCGTGGCGCGCGCGCGCGAGTGGCGGACCGCGCGCGCGCATGCGCGGC 732
QY      706 GCCAAGCAGACCTTGAAGAGCAGATGCGCTGATGACCAAGACCGCGCGCGCGCGTGG 765
Db      733 ACCAAGCAGACCTTGAAGAGCAGATGCGCTGATGACCAAGACCGCGCGCGCGCGTGG 792
QY      766 GCGGACATCTACAGCGGTGATCATCTGGGCTTGAACAGATGTGTGCGATGTACAGC 825
Db      793 GCGGACATCTACAGCGGTGATCATCTGGGCTTGAACAGATGTGTGCGATGTACAGC 852
QY      826 CCCGTGAGCATCTTGAAGATCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 885
Db      853 CCCAAGCAGATCTTGAAGATCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 912
QY      886 CGCTTCTCAAGACCTTGGCGGCCGAGCAGGCCCAACCAAGACGTGAAGAACTGGATGACC 945
Db      913 CGCTTCTCAAGACCTTGGCGGCCGAGCAGGCCCAACCAAGACGTGAAGAACTGGATGACC 972
QY      946 GAGACCTGCTGTGACAGAAAGCCCAACCCCGACTGCAAGACCATCTGCGCGCTCTGGGC 1005
```

Db 973 GAGACCTGCTGTGTCAGAAACCCCGACTGCAAGACCATCTGAAGCTCTCGGC 1032

QY 1006 CCGCGCGCCACCTTGAGAGATGATGACCGCTGCCAGGCGTGGCGCGCCCGGCAC 1065

Db 1033 CCGCGCGCCACCTTGAGAGATGATGACCGCTGCCAGGCGTGGCGCGCCCGGCAC 1092

QY 1066 AAGCGCGCGCTGCTGGCCGAGCGATGAGCGAGG---CCAAGCGTGAACATCATGATG 1122

Db 1093 AAGCGCGCGCTGCTGGCCGAGCGATGAGCGAGCGAGCGCGGACCATCATGATG 1152

QY 1123 CAGAAGCAACTTCAAGGCGCCCGCGCAAGTCAAGTCTTCAACTGCGGCAAGAG 1182

Db 1153 CAGCGCGCAACTTCCGCAACAGCGGAAACGTCAGTCTTCAACTGCGGCAAGAG 1212

QY 1183 GCGCACATGCGCAAGAACTGCGCGCGCGCGCAAGAGGCTGCTGGAAGTGGCAAG 1242

Db 1213 GCGCACACCGCGAGAACTGCGCGCGCGCGCAAGAGGCTGCTGCGCGCGCGC 1272

QY 1243 GAGGCGCACAGATGAAGACTGCAACCGAGCGCGCAACTTCTGCGCAAGATCTGG 1302

Db 1273 GAAGCACCAAAATGAAGATTGCACTGAGAGACAGCTAAATTTTAAAGGAAGATCTGG 1332

QY 1303 CCGAGCCACAAGGCGCGCGCGCAACTTCTGCAAGACCGAGCGAGCGCGCGCGCC 1362

Db 1333 CCTTCTTCAAGGAGGAGCGCGAGGCAATTTCTTCAAGAGCAG-----A 1374

QY 1363 ACCGTGCGCACCGCGCGCGCGCGCGAGAGCTTCCGCTTGAAGAGACACCGCGCGCC 1422

Db 1375 CCAGAGCCAAAGCGCGCGCGCGCGAGAGAGCTTCAAGTTGGGAGAGAGAAACAACTGCC 1434

QY 1423 AAGCAGAGCCCAAGAGAGCGCGAGCGCTTACCGAGCGCGCGCGCGCGCGCGCTG 1482

Db 1435 TCTCAGAGCAGAGCGCGAGCGAGCTGATCTTAACTTCCCTCAGATCACTC 1494

QY 1483 TTGCGCAGCGCGCGCGCGCGCGAGTAA 1509

Db 1495 TTTGCGCAGCAGCGCGCGCGCGCGAGTAA 1521

RESULT 9  
US-09-475-515-79  
Sequence 79, Application US/09475515A

GENERAL INFORMATION:  
PATENT No. 6602705  
APPLICANT: BARNETT, Susan  
APPLICANT: ZUR MEGEDE, Jan  
APPLICANT: SRIVASTAVA, Indresh  
APPLICANT: LIAN, Ying  
APPLICANT: HARTOG, Karin  
APPLICANT: LIU, Hong  
APPLICANT: GRIBER, Catherine  
APPLICANT: SELBY, Mark  
APPLICANT: WALKER, Christopher  
TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION  
TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES  
FILE REFERENCE: 1621.002  
CURRENT APPLICATION NUMBER: US/09/475,515A  
CURRENT FILING DATE: 1999-12-30  
NUMBER OF SEQ ID NOS: 90  
SOFTWARE: Patent Ver. 2.0  
SEQ ID NO 79  
LENGTH: 1865  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: GP2  
US-09-475-515-79

Query Match 74.1%; Score 1117.8; DB 4; Length 1865;  
Best Local Similarity 85.1%; Pred. No. 1.8e-157;  
Matches 1299; Conservative 0; Mismatches 192; Indels 36; Gaps 3;  
QY 1 ATGGGCGCGCGCGCAGCATCTGCGCGCGAGAACTGACAAAGTGAGAGAAATCCGC 60

Db 13 ATGGGCGCGCGCGCAGCGCTGAGCGCGCGAGCTGACAAAGTGAGAGAAATCCGC 72

QY 61 CTGCGCGCGCGCGCAAGAGCACTAATGCTGAAGCACTTGCTGGCCAGCGCGAG 120

Db 73 CTGCGCGCGCGCGCAAGAGCACTAATGCTGAAGCACTTGCTGGCCAGCGCGAG 132

QY 121 CTGAGAGCTTCCCTGAACCCCGGCTGCTGAGAGACCGCGAGGGCTGCAAGCATG 180

Db 133 CTGAGAGCTTCCCTGAACCCCGGCTGCTGAGAGACCGCGAGGGCTGCGCGCATG 192

QY 181 ATGAAGCAGCTGAGCGCGCGCTGCAAGCGCGCAAGAGAGCTGCGCAGCTGTACAAC 240

Db 193 CTGCGCAGCTGCAAGCGCGCGCTGCAAGCGCGCAAGAGAGCTGCGCAGCTGTACAAC 252

QY 241 ACCGTGCGCACCTGTACTGCTGTCAGCGCGCGCATGAGTCCGCGCAACCAAGAGGCC 300

Db 253 ACCGTGCGCACCTGTACTGCTGTCAGCGCGCATGAGTCCGCGCAACCAAGAGGCC 312

QY 301 CTGACAAAGATGAGAGAGAGAGCAAGATCCAGCAAGAGACCGAGCGCAAGAGAG 360

Db 313 CTGAGAGAGATGAGAGAGAGAGCAAGATCCAGCAAGAGAGCGCGCGCGCGC 372

QY 361 GCGCAGCGCA-----AGGTGAGCGCAAGTACCTATCGTGCAGAACTG 405

Db 373 GCGCGCGCGCACCGCGCAACAGCAGCGAGTGAACCAAGTACCTATCGTGCAGAACTG 432

QY 406 CAGGCGCAGATGCTGACAGGCGCATCAGCGCGCGCACCTGAAAGCTGGAAGTG 465

Db 433 CAGGCGCAGATGCTGACAGGCGCATCAGCGCGCGCACCTGAAAGCTGGAAGTG 492

QY 466 ATCGAGAGAGGCGCTTACGCGCGCGAGGTATCCCATGTTCACCGCGCTGAGCGAGGGC 525

Db 493 GTGAGAGAGAGGCGCTTACGCGCGCGAGGTATCCCATGTTCAGCGCGCTGAGCGAGGGC 552

QY 526 GCCACCGCGCGAGACTGAACAGATGTTGAACCGTGGCGCGCACAGCGCGCATG 585

Db 553 GCCACCGCGCGAGACTGAACAGATGTTGAACCGTGGCGCGCACAGCGCGCATG 612

QY 586 CAGATGCTGAAGAGACCATCAAGAGAGCGCGCGAGTGGAGCGCGTGCACCGCGTG 645

Db 613 CAGATGCTGAAGAGAGACCATCAAGAGAGCGCGCGAGTGGAGCGCGTGCACCGCGTG 672

QY 646 CAGCGCGCGCGCGTGGCGCGCGCGCGAGATGCGCGAGCGCGCGCGCGCATGCGCGGC 705

Db 673 CAGCGCGCGCGCGTGGCGCGCGCGCGAGATGCGCGAGCGCGCGCGCGCATGCGCGGC 732

QY 706 GCCACGAGACCGCTGAGAGAGAGATGCGCTGATGACCAAGACCGCGCGCGCGTG 765

Db 733 ACCACGAGACCGCTGAGAGAGAGATGCGCTGATGACCAAGACCGCGCGCGTG 792

QY 766 GCGGACATCTACAAGCGGTGATCATCTGGCGCTGAACAAGATGTCGAGATGACAGC 825

Db 793 GCGGAGATCTACAAGCGGTGATCATCTGGCGCTGAACAAGATGTCGAGATGACAGC 852

QY 826 CCGGTGAGCATCTGAGCATTCGCGCGAGGCGCGCAAGAGCGCTTCCGCGAGTACGTGAGC 885

Db 853 CCGGTGAGCATCTGAGCATTCGCGCGAGGCGCGCAAGAGCGCTTCCGCGAGTACGTGAGC 912

QY 886 CGCTTCTTCAAGACCGTGGCGCGCGAGCGCGCAAGAGCGTGAAGATGATGAGC 945

Db 913 CGCTTCTTCAAGACCGTGGCGCGCGAGCGCGCAAGAGCGTGAAGATGATGAGC 972

QY 946 GAGACCTGCTGTGTCAGAACCGCAACCGCGCACTGCAAGACCAATCTGAGGCTCTCGGC 1005

Db 973 GAGACCTGCTGTGTCAGAACCGCAACCGCGCACTGCAAGACCAATCTGAGGCTCTCGGC 1032

QY 1006 CCGCGCGCGCACCTTGAGAGATGATGACCGCTGCGCAAGGCGTGGCGCGCGCGCAC 1065

Db 1033 CCGCGCGCGCACCTTGAGAGATGATGACCGCTGCGCAAGGCGTGGCGCGCGCGCAC 1092

QY 1066 AAGCGCGCGCTGCTGGCCGAGCGATGAGCGAGG---CCAAGCGTGAACATCATGATG 1122



Db	1093	AMGCCCCGTGCTGCGCCGAGCGATGAGCCAGGTGACGAACCCCGCGACCATCAATGATG	1152
Qy	1123	CAGAAGAGCAACTTCAMGGGCCCCCGCGCAACGTCAGTGTCTTCAACTGCGGCAAGAG	1182
Db	1153	CAGCGCGCAACTTCGCAACCAAGCGGAAAGACCGTCAGTGTCTTCACTGCGGCAAGAG	1212
Qy	1183	GGCCACATGCGCAAGAACTGCGCGCCCCCGGCAAGAAAGGCTGCTGGAAGTGCAGGCAAG	1242
Db	1213	GGCCACACCGCGAGAACTGCGCGCCCCCGGCAAGAAAGGCTGCTGCGCTGCGCGCGC	1272
Qy	1243	GAGGCGCACCATGAAAGCACTGCAACGAGCGCCAGGCCAACTTCTGCGGCAAGATCTGG	1302
Db	1273	GAAAGACCAACAATGAAAGATTGCACTGAGAGACAGGCTAATTTTAAAGGGAAGATCTGG	1332
Qy	1303	CCCAAGCCAAAGGGCGCGCCCGSCAACTTCTGTGAGAAACCGCAGCGAGCCCGCGCCCC	1362
Db	1333	CCTTCTTAACAAGGGAAGGCGCAGGGAATTTCTTTCAGAGCAG-----A	1374
Qy	1363	ACCGTGCCCAACCGCCCCCGCGCCGAGAGCTTCCGCTTCAGAGAGACCAACCCCGCCCC	1422
Db	1375	CCAGAGCCAAACAGCCCCACCAAGAGAGAGCTTCAGGTTTGGGGAGAGAAACCAACTCCC	1434
Qy	1423	AAGCAGAGGCCCAAGGACCGGAGCCCTTACCGCGAGCCCCCTGACCGCCCTGCGCAGCTTG	1482
Db	1435	TCTCAGAGCAGGAGGCCGATAGCAAGGAACGTGTATCCTTAACTTCCCTCAGATCACTTC	1494
Qy	1483	TTCCGCGAGCGCCCCCTTGAGCCAGTTAA	1509
Db	1495	TTTGGCAACGACCCCTCGTCACAGTTAA	1521

## RESULT 10

```

US-09-475-515-6
; Sequence 6, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 4319
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: HIV-Gag-polymerase
US-09-475-515-6

```

Query Match	74.1%;	Score 1117.8;	DB 4;	Length 4319;
Best Local Similarity	85.1%;	Pred. No. 1.8e-157;		
Matches 1299; Conservative	0;	Mismatches 192;	Indels 36;	Gaps 3;

OY	1	ATGGGCGCCCGCGCCAGCATCTTCGCGCGCGGAGAGCTGGACAACTGGGAGAAATCCGC	60
Db	7	ATGGGCGCCCGCGCCAGCGTGTCTAGCGGCGGCGGAGCTGGACAACTGGGAGAAATCCGC	66
OY	61	CTGCGCCCCGCGCGCAAGAGCATACATGCTGAAGCACTGGTGTGGCCAGCCGCGAG	120
Db	67	CTGCGCCCCGCGCGCAAGAGATACAACTGTGAAGCACATCGTGTGGCCAGCCGCGAG	126

QY	121	CTGAGGGCTTGGCCCTGTAACCCCGGCTGCTGAGACCGCCGAGGGCTGCAAGCAGATC	180
Db	127	CTGAGCGCTTGGCGGTGAACCCCGGCTGCTGAGACCAAGAGGGCTGCGCCAGATC	186
QY	181	ATGAAGCAGCTGCAGCCCGCTGCAAGACCGGCACCCGAGAGCTGCGCAGCTGTACAAC	240
Db	187	CTGGCCAGCTGCAGCCCGCTGCAAGACCGGCACCGAGAGCTGCGCAGCTGTACAAC	246
QY	241	ACCGTGGCCAACCTGTACTGCTGTCACCGCGCATTCAGCTCAAGGACCAAGAGGCC	300
Db	247	ACCGTGGCCAACCTGTACTGCTGTCACCGCGCATTCAGCTCAAGGACCAAGAGGCC	306
QY	301	CTGACAAGAATCGAGAGGAGCAGAACAAATGCCAGCAGAAAGACCAGCAGGCCCAAGAG	360
Db	307	CTGAGAGAAGATCGAGAGGAGCAGAACAAATGCCAGAGAGGCCAGCAGGCCCGCGCC	366
QY	361	GCCGACGGCA-----AGGTAGCCAGAACTAACCCCATCGTGCAGAACTTG	405
Db	367	GCCGCGCGCAACCGCAACAGCAGCCAGGTGAGCCAGAACTAACCCCATCGTGCAGAACTTG	426
QY	406	CAGGGCCAGATGCTGCAACAGGCCATCAGCCCCCGCACCTGAACGCTGGGTGAAGTG	465
Db	427	CAGGGCCAGATGCTGCAACAGGCCATCAGCCCCCGCACCTGAACGCTGGGTGAAGTG	486
QY	466	ATCGAGAGAAGGCTTCAAGCCCGAGGTGATCCCATGTTCAACGCGCTGAGCGAGGGC	525
Db	487	GTGAGAGAAGAGGCTTCAAGCCCGAGGTGATCCCATGTTCAAGCGCTGAGCGAGGGC	546
QY	526	GCCACCCCGCAGAGCTGAAACGATGTTGAACACCGTGGCGGCCACACAGCCGCCATG	585
Db	547	GCCACCCCGCAGAGCTGAAACGATGTTGAACACCGTGGCGGCCACACAGCCGCCATG	606
QY	586	CAGATGCTGAAGACACCATCAACGAGAGGGCGCCGAGTGGGACCGCTGCACCCCGTG	645
Db	607	CAGATGCTGAAGAGACCATCAACGAGAGGGCGCCGAGTGGGACCGCTGCACCCCGTG	666
QY	646	CAGGCGGGCCGCTGGCCCCCGGCAGATGCGGACCCCGCGGACGACATGCGCGGC	705
Db	667	CAGGCGGGCCCATGCGCCCCCGGCAGATGCGGAGCCCGCGGACGACATGCGCGGC	726
QY	706	GCCACCAAGACCTGCAAGAGCAGATGCTGATGAACCAACCCCGCTGCGCCGTG	765
Db	727	ACCAACCAAGACCTGCAAGAGCAGATGCTGATGAACCAACCCCGCATGCCGTG	786
QY	766	GCGCAGATCTACAAGCGGTGATCATCTGGGCTGAACAAGATGTGCGATGTACAGC	825
Db	787	GCGCAGATCTACAAGCGGTGATCATCTGGGCTGAACAAGATGTGCGATGTACAGC	846
QY	826	CCCGTGAGCAATCTGCAATCCGCAAGGCCCAAGAGCCCTTCCGCACTACGTGAC	885
Db	847	CCCAACAGCAATCTGCAATCCGCAAGGCCCAAGAGCCCTTCCGCACTACGTGAC	906
QY	886	CGCTTCTTCAAGACCTTGGCGCGCGAGCAGGCCACCCAGACGTGAAGAACTGATGACC	945
Db	907	CGCTTCTTCAAGACCTTGGCGCGCTGAGCAGGCCAGCCAGACGTGAAGAACTGATGACC	966
QY	946	GAGACCTTGTGCTGCAAGCGCAACCCGACTGCAAGACCATCTGCGGCTCTCGGC	1005
Db	967	GAGACCTTGTGCTGCAAGCGCAACCCGACTGCAAGACCATCTGGAAGCTCTCGGC	1026
QY	1006	CCCGCGGCAACCTGAGAGATGATGACCGCTGCAGGGCGTGGGCGGCCCGGCCAC	1065
Db	1027	CCCGCGGCAACCTGAGAGATGATGACCGCTGCAGGGCGTGGGCGGCCCGGCCAC	1086
QY	1066	AAGGCCCGCTGTGGCCGAGGCGATGAGCCAGTGACCAACCCGCGCAGCATCATGATG	1122
Db	1087	AAGGCCCGCTGTGGCCGAGGCGATGAGCCAGTGACCAACCCGCGCAGCATCATGATG	1146
QY	1123	CAGAAAGCACTTCAAGGGCCCCCGCGCAACGTCAAGTCTTCACTGCGGCAAGAG	1182
Db	1147	CAGCGCGCACTTCCGCAACCAAGCGAAGACGTCAAGTCTTCACTGCGGCAAGAG	1206
QY	1183	GCGCAGATCGCCAAAGACTGCGCGCCCCCGCAAGAGGCTGCTGAAGTGCAGAG	1242



```
Db 1207 GGGCCACCGCCAGGAATGCGCGCGCCCGCCGCAAGAGGCTGCTGCGCTGCGCGC 1266
Qy 1243 GAGGGCCACAGATGAAGACTGCAACCGAGCGCCCACTTCTGGGCAAGATCTGG 1302
Db 1267 GAAGGACCAATATGAAGATTGCACTGAGAGACAGGCTAATTTTATGGAAGATCTGG 1326
Qy 1303 CCCAGCCCAAGAGGCGCGCCCGCACTTCTGCAAGACCGCAGCGCGCGCGCC 1362
Db 1327 CTTCTTCAAGAGGGAAGGCGCAATTTTCTCAGAGCAG-----A 1368
Qy 1363 ACCGTGCCCCACCGCGCGCGCGCGAGCTTCCGCTTCAAGAGACCAACCGCGCGCC 1422
Db 1369 CCAGAGCCCAACAGCCCGCCAGAGAGAGCTTCAAGTTTGGGAGAGAAACAATCTCC 1428
Qy 1423 AAGCAGAGCCCAAGAGACCGCGAGCCCTTACCGGAGCCCTGACCGCGCGAGCTG 1482
Db 1429 TCTCAGAGAGAGAGAGCGCGATGACAGAGACTGTATCTTTAATCTTCCCTCAGATCACTC 1488
Qy 1483 TTGGGAGCGCGCGCGCGCGCTGAGCCAGTAA 1509
Db 1489 TTGGGAGAGAGAGCGCGCGCTGCTCAGAGTAA 1515
```

## RESULT 11

```
US-09-475-515-9
; Sequence 9, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GRIBER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475, 515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 9
; LENGTH: 1268
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic Gag
; OTHER INFORMATION: common region
US-09-475-515-9
```

Query Match 69.5%; Score 1049.2; DB 4; Length 1268;

Best Local Similarity 90.8%; Pred. No. 2.5e-147;

Matches 1146; Conservative 0; Mismatches 98; Indels 18; Gaps 2;

```
Qy 1 AATGGGCGCGCGCCAGCATCTGCGCGCGAGAGCTGACAAGTGGAGAAGATCCGC 60
Db 7 AATGGGCGCGCGCGCGCGCTGTGAGCGCGCGAGCTGACAAGTGGAGAAGATCCGC 66
Qy 61 CTGGCGCGCGCGCGCGCAAGAGCACTAATGTGAAGCACTGTGTGGCGCAGCGCGAG 120
Db 67 CTGGCGCGCGCGCGCGCAAGAGCAAGTGAAGCACTGTGTGGCGCAGCGCGAG 126
Qy 121 CTGAGAGGCTTCTGCGCTGAACCCCGCGCTGTGAGACCGCGAGGCTGCAAGCAGATC 180
Db 127 CTGAGAGGCTTCTGCGCTGAACCCCGCGCTGTGAGACCGCGAGGCTGCGCAGATC 186
Qy 181 ATGAAGCAGCTGACCGCGCGCTGCAAGCGGACCGAGAGCTGCGAGCTGTACAC 240
Db 187 CTGGGCGCAGCTGACCGCGCGCTGCAAGCGGACCGAGAGAGCTGCGAGCTGTACAC 246
```

```
Qy 241 ACCGTGACCACTCTGTACTGCTGACACGCGCGCATGAGTCCGCGACCAAGAGGCC 300
Db 247 ACCGTGACCACTCTGTACTGCTGACACGCGCATGAGTCCGCGACCAAGAGGCC 306
Qy 301 CTGACAAAGATCGAGAGAGAGCAAGAACTCCAGCAGAAAGACCAGAGGCCAAGAG 360
Db 307 CTGAGAAAGATCGAGAGAGAGCAAGAACTCCAGAGAAAGGCCAGAGGCCCGCC 366
Qy 361 GCCAGCGCA-----AGGTAGCCAGAACTACCCCATCTGTGCAAACTTG 405
Db 367 GCCCGCGCAACCGGCAACAGCAGCCAGGTAGCCAGAACTACCCCATCTGTGCAAACTTG 426
Qy 406 CAGGGCCAGATGCTGCAACAGGCCATCAGCCCGCACCTGAAAGCTGGTGAAGTG 465
Db 427 CAGGGCCAGATGCTGCAACAGGCCATCAGCCCGCACCTGAAAGCTGGTGAAGTG 486
Qy 466 ATCGAGAGAGAGGCTTCAAGCCCGAGGTGATCCCATGTTCAACCGCGCTGAGCGAGGC 525
Db 487 GTGAGAGAGAGAGGCTTCAAGCCCGAGGTGATCCCATGTTCAAGCGCGCTGAGCGAGGC 546
Qy 526 GCCACCGCCAGAGAGCTGAAACAGATGTTGAACACCGTGGCGGCCACAGCGCCCATG 585
Db 547 GCCACCGCCAGAGAGCTGAAACAGATGTTGAACACCGTGGCGGCCACAGCGCCCATG 606
Qy 586 CAGATGCTGAAGAGCAACCACTCAACGAGAGCGCGCCGAGTGGAGCCGCTGACCCCGTG 645
Db 607 CAGATGCTGAAGAGCAACCACTCAACGAGAGCGCGCCGAGTGGAGCCGCTGACCCCGTG 666
Qy 646 CAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 705
Db 667 CAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 726
Qy 706 GCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 765
Db 727 ACCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 786
Qy 766 GCGGACATCTCAAGAGCGGTGATCATCTGCGCGCTGAAACAAGATGTCGATGACAGC 825
Db 787 GCGGAGATCTCAAGAGCGGTGATCATCTGCGCGCTGAAACAAGATGTCGATGACAGC 846
Qy 826 CCGGTGAGATCTGAGACATCCGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 885
Db 847 CCCACAGAGATCTGAGACATCCGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 906
Qy 886 CGCTTCTTCAAGAGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 945
Db 907 CGCTTCTTCAAGAGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 966
Qy 946 GAGACCTGCTGTGTGCAAGCGCAACCGCGAGCTGCAAGACCATCTGCGCGCTCTCGGC 1005
Db 967 GAGACCTGCTGTGTGCAAGCGCAACCGCGAGCTGCAAGACCATCTGAGAGCTCTCGGC 1026
Qy 1006 CCGGCGCGCAACCTGAGAGAGATGATACCGCTGCGCGAGGCGGTGGCGCGCGCGCGCAC 1065
Db 1027 CCGGCGCGCAACCTGAGAGAGATGATACCGCTGCGCGAGGCGGTGGCGCGCGCGCGCAC 1086
Qy 1066 AAGGCGCGGTCTGCGCGAGGCGGATGAGCGAG--CCAACAAGCTGAACATCATGATG 1122
Db 1087 AAGGCGCGGTCTGCGCGAGGCGGATGAGCGAGCTGAACCAACCGCGAGCATCATGATG 1146
Qy 1123 CAGAGAGCAACTTCAAGGCGCGCGCGCGCAACGTCAAGTGTCTTCAACTGCGCGCAAGAG 1182
Db 1147 CAGCGCGCAACTTCCGCAACCAAGCGAGACCGTCAAGTGTCTTCAACTGCGCGCAAGAG 1206
Qy 1183 GCGCACATGCGCAAGAACTGCGCGCGCGCGCGCAAGAGAGGCTGTGAAAGTGGCGAAG 1242
Db 1207 GCGCACATGCGCAAGAACTGCGCGCGCGCGCGCAAGAGAGGCTGTGCGCGTGGCGCGC 1266
Qy 1243 GA 1244
Db 1267 GA 1268
```

RESULT 12  
US-09-552-950-2  
Sequence 2, Application US/09552950  
Patent No. 6541248  
GENERAL INFORMATION:  
APPLICANT: Oxford Biomedica (UK) Limited  
TITLE OF INVENTION: Anti-Viral Vectors  
FILE REFERENCE: 674524-2004  
CURRENT APPLICATION NUMBER: US/09/552,950  
CURRENT FILING DATE: 2000-04-20  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 4307  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: gagpol-SYNsp - codon  
US-09-552-950-2

Query Match 61.6%; Score 929.4; DB 4; Length 4307;  
Best Local Similarity 77.7%; Pred. No. 1.2e-129;  
Matches 1173; Conservative 0; Mismatches 306; Indels 30; Gaps 3;

QY 1 ATGGGCGCCGCGCGCAGCATCTGCGCGCGGAGAGAGCTGAGCAATGGGAGATCCGC 60  
DB 1 ATGGGCGCCGCGCGCAGCATCTGCGCGCGGAGAGAGCTGAGCAATGGGAGATCCGC 60  
QY 61 CTGCGCGCCGCGCGCAGAGAGCACTACATGCTGAAGCACTGTGTGGGCGCAGCGCGAG 120  
DB 61 CTGCGCGCCGCGCGCAGAGAGCACTACATGCTGAAGCACTGTGTGGGCGCAGCGAG 120  
QY 121 CTGAGGCGCTTGGCCCTGAAACCCCGGCTGTGAGAGACCGCGAGGGCTGCAAGCAATC 180  
DB 121 CTGAGGCGCTTGGCCCTGAAACCCCGGCTGTGAGAGACCGCGAGGGCTGCAAGATC 180  
QY 181 ATGAAGCAGCTGAGCCCGCGCTGAGACCGCGAGCGGAGCTGCGCAGCTGTACAC 240  
DB 181 CTGCGCGCACTGAGCCCGCGCTGAGACCGCGAGCGGAGCTGCGCAGCTGTACAC 240  
QY 241 ACCGTGGCCACCTGTACTGCTGCTGAGACCGCGAGCTGAGGTCGCGCAACCAAGAGGCC 300  
DB 241 ACCGTGGCCACCTGTACTGCTGCTGAGACCGCGAGCTGAGGTCGCGCAACCAAGAGGCC 300  
QY 301 CTGAGCAAGATCGAGGAGAGAGAGCAAGTCCAGAGAGAGCCAGAGGCCAAGAGAG 360  
DB 301 CTGATTAATTCAGAGAGAGAGAGATTAAGAGAGAGAGAGAGAGGCCAGAGGCC 360  
QY 361 GCCGACGG-----CAAGGTAGCCAGAGAGTAACTACCCCATCTGAGAGAGAGGGGC 411  
DB 361 GACACCGGAGCAGAGACAGAGTCAAGCAAGTAACTACCCCATCTGAGAGAGAGGGGC 420  
QY 412 CAGATGTGCAACGAGGCGCATGAGCCCGCGCAGCTGAAAGCTGTGGTGAAGTGTGAG 471  
DB 421 CAGATGTGCAACGAGGCGCATGAGCCCGCGCAGCTGAAAGCTGTGGTGAAGTGTGAA 480  
QY 472 GAGAGGCGCTTCAAGCCCGAGGTGATCCCATGTTCACCGCTGAGCGAGGGCGCACCC 531  
DB 481 GAGAGGCGCTTCAAGCCCGAGGTGATCCCATGTTCACCGCTGAGCGAGGGCGCACCC 540  
QY 532 CCCAGAGACCTGAACACGATGTGAACACCGTGGGCGCGCAGCAGAGCGCGCATGAGATG 591  
DB 541 CCCAAGATCTGAACACGATGTTCACACAGTGGGCGCGCAGCAGAGCGCGCATGAGATG 600  
QY 592 CTGAAGAGACCATCAACGAGAGAGCGCGCAGTGGAGCCGCTGCAACCCCGTGAAGGCC 651  
DB 601 CTGAAGAGACCATCAATGAGAGAGCGCTGCGCAATGGATGTGTGATCCCGTGAAGCGCA 660  
QY 652 GGGCCGCTGGCGCGCGCGCAGAGCGGAGCCCGCGCGCAGAGCATGCGCGCGCGCACCC 711  
DB 661 GGGCCCATGCGACCGGGCGCAGATGCGTGAAGCAAGGGGCTCAGACATGCGCGCGAGCAT 720

QY 712 AGCAACCTGAGAGAGAGATTCGCTGTGATGACCAAGCAACCCCGGTGCGCGGGCGAC 771  
DB 721 AGTACCTTCAGAGAACAGATTCGGCTGTGATGACCAACCAACCCCATCCCGGTGGAGAA 780  
QY 772 ATCTACAGCGGTGATCATCTGCGCTGGAACAAGATCTGTGCGGATGTACAGCCCGT 831  
DB 781 ATCTACAAAGCTGTGATCATCTGCGCTGGAACAAGATCTGTGCGGATGTATAGCCCTAC 840  
QY 832 AGCATCTGAGCATTCGCGCGAGGGCGCGCAAGAGCCCTTCCGCGACTACGTGAGCCGCTTC 891  
DB 841 AGCATCTGAGCATTCGCGCGAGGGCGCGCAAGAGCCCTTCCGCGACTACGTGAGCCGCTTC 900  
QY 892 TTCAAGACCTGCGCGCGAGAGAGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 951  
DB 901 TACAAAGAGCTCCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960  
QY 952 CTGCTGTGAG 1011  
DB 961 CTGCTGTGAG 1020  
QY 1012 GCCACCTGAGAGAGAGATGATGACCGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAG 1071  
DB 1021 GCTACCTGAGAGAGAGATGATGACCGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080  
QY 1072 CGCGTCTGCGCGAG 1128  
DB 1081 CGCGTCTGCGCGAG 1140  
QY 1129 AGCAACTTCAAG 1188  
DB 1141 GGCACCTTCAAG 1200  
QY 1189 ATGCGCAAG 1248  
DB 1201 ACAGCCCGCAACTGAG 1260  
QY 1249 CACAGATGAAG 1308  
DB 1261 CACAGATGAAG 1320  
QY 1309 CACAAG 1368  
DB 1321 TACAAG 1362  
QY 1369 CCCACGCGCGCGCGCGAG 1428  
DB 1363 CCAAG 1422  
QY 1429 GAGCCCAAG 1488  
DB 1423 AAGCAG 1482  
QY 1489 AGCGGCGCGC 1497  
DB 1483 AGCGGCGCGC 1491

RESULT 13  
US-09-552-950-5  
Sequence 5, Application US/09552950  
Patent No. 6541248  
GENERAL INFORMATION:  
APPLICANT: Oxford Biomedica (UK) Limited  
TITLE OF INVENTION: Anti-Viral Vectors  
FILE REFERENCE: 674524-2004  
CURRENT APPLICATION NUMBER: US/09/552,950  
CURRENT FILING DATE: 2000-04-20  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 5  
LENGTH: 9772  
TYPE: DNA



```
QY 1 ATGGGCGCCCGCCGACATCTCTGCGCGGAGAGCTGCACTAGTGGAGAGATCCG 60
DB 770 ATGGGTGGAGAGCGTCACTATTAAAGCGGGGAGAAATTAGATCGATGGGAAAAATTCCG 829
QY 61 CTGCGCCCGCGCGGAGAGCACTACATCTGAAGCACCCTGCTGCGGCGACCGCGAG 120
DB 830 TTAAAGCCAGGGGAGAAAGAAAGTAAAGCAATCGTATGCGCAAGCAGGGAG 889
QY 121 CTGAGGCGCTTCCCTGAAACCCCGGCTGAGAGACCGCGAGGCTGCAAGCAGATC 180
DB 890 CTAGAACGATTCGAGTAAATCTGCGCTGTAGAAACATCAGAAAGCTGTAGACAAATA 949
QY 181 ATGAAGCACTGCAAGCCCGCTGCAAGCCGCAAGAGAGCTGCGAGCTGTACAAC 240
DB 950 CTGGAGCACTACAACTATCTTCAAGACAGATCAGAGAGCTTCATCACTATACAAAC 1009
QY 241 ACCGTGCGCACTCTGTAATCTGCTGCAAGCCGCGCATCGAGTCCGCAACCAAGAGGCG 300
DB 1010 ACAGTAGCAACCTCTATTTGTGTGCAAGCGGATCGAGATCAAGACCAAGAGAGCT 1069
QY 301 CTGACAAAGATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 353
DB 1070 TTAGACAAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1129
QY 354 --CAAGAGCGCGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 411
DB 1130 GACACAGAGACAGCAATCAGGTCAAGCAAAATTAACCTATAGTGCAGAGACATCCAGGG 1189
QY 412 CAGATGTGCAACAGGACATCAAGCCCGCAAGCTGAAAGCTGAGTGAAGTATGAG 471
DB 1190 CAATGTGATCATGAGGCAATCACTAGAACTTTAAATGCAATGGGTAAAGTATAGAA 1249
QY 472 GAGAGAGCTTCAAGCCCGAGAGATCCCATGTTCACCGCTGAGCGAGGCGCGAC 531
DB 1250 GAGAGAGCTTCAAGCCCGAGAGATCCCATGTTCACCGCTGAGCGAGGCGCGAC 1309
QY 532 CCGCAGAGCTGAAACAGATGTGAACACCGTGGGCGGCGCACAGGCGCGCATGCAAGT 591
DB 1310 CCACAGAGCTGAAACAGATGTGAACACCGTGGGCGGCGCACAGGCGCGCATGCAAGT 1369
QY 592 CTGAAGAGACCATCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 651
DB 1370 TTAAAGAGACCATCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1429
QY 652 GCGCCCGTGGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 711
DB 1430 GCGCTATTCACAGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1489
QY 712 AGCACCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 771
DB 1490 AGTACCCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1549
QY 772 ATCTACAAGCGGTGATCATCTGCGGCTGAAACAGATCGTGGAGTATACGCCGTG 831
DB 1550 ATCTACAAGAGGTGATCATCTGCGGCTGAAACAGATCGTGGAGTATACGCCGTAC 1609
QY 832 AGCATCTGAGATCCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 891
DB 1610 AGCATCTGAGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1669
QY 892 TTCAAGAGACCTGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 951
DB 1670 TATAAAACCTTAAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1729
QY 952 CTGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1011
DB 1730 TTGTGTGTCAGAAATGCGAACTGATGTAGACATCTGAAAGCTCTGCGCCAGCG 1789
QY 1012 GCCACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1071
DB 1790 GCTAACACTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1849
QY 1072 CGCGTGTGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1128
```

```
DB 1850 AGAGTTTGGCCGAGCGGATGAGCCAGGTCAGAACTCGCGGACCATATATGTCAGAGA 1909
QY 1129 AGCAACTTCAAGGGCCCCCGCGCAACGTCAAGTCTTCAACTGCGGCAAGAGAGGCCAC 1188
DB 1910 GGCACCTTCGGAACACAGCGAAGATGTCAGTCTTCAATTGTGCAAGAGAGGCCAC 1969
QY 1189 ATGCGCAAGACTGCGCGCGCGCGCGCAAGAGGCTGTGAAAGTGTGCGCAAGAGAGG 1248
DB 1970 ACCGCCAGAGACTGCGCGCGCGCGCGCAAGAGGCTGTGAAAGTGTGAAAGAGAG 2029
QY 1249 CACCAATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1308
DB 2030 CACCAATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2089
QY 1309 CACAAGGCGCGCGCGCGCACTTCTGTCAGAAACCGCAGAGAGAGAGAGAGAGAG 1368
DB 2090 TACAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2131
QY 1369 CCCACGCGCGCGCGCGCGAGAGCTTCCGCTTGCAGAGAGAGAGAGAGAGAGAGAG 1428
DB 2132 CCAACAGCGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2191
QY 1429 GAGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1488
DB 2192 AAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2251
QY 1489 AGCGCGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1509
DB 2252 AAGCAGCGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2272
```

```
RESULT 15
US-09-184-418C-82
; Sequence 82, Application US/09184418C
; Patent No. 6492110
; GENERAL INFORMATION:
; APPLICANT: Hahn, Beatrice
; APPLICANT: Gao, Peng
; APPLICANT: Shaw, George
; TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN
; TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS TYPE 1
; FILE REFERENCE: D6287
; CURRENT APPLICATION NUMBER: US/09/184,418C
; CURRENT FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 112
; SEQ ID NO 82
; LENGTH: 1496
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; OTHER INFORMATION: isolate=96ZM751.3; gene=gag
US-09-184-418C-82
```

```
Query Match 48.7%; Score 734.2; DB 4; Length 1496;
Best Local Similarity 69.0%; Pred. No. 8.5e-101;
Matches 1020; Conservative 0; Mismatches 458; Indels 1; Gaps 1;
QY 1 ATGGGCGCCCGCCGACATCTCTGCGCGGAGAGCTGCACTAGTGGAGAGATCCG 60
DB 1 ATGGGTGGAGAGCGTCACTATTAAAGCGGGGAGAAATTAGATCGATGGGAAAAATTCCG 60
QY 61 CTGCGCCCGCGCGGAGAGCACTACATCTGAAGCACCCTGCTGCGGCGACCGCGAG 120
DB 61 CTGAGGCGCTTCCCTGAAACCCCGGCTGAGAGACCGCGAGGCTGCAAGCAGATC 180
QY 121 CTGAGGCGCTTCCCTGAAACCCCGGCTGAGAGACCGCGAGGCTGCAAGCAGATC 180
DB 121 CTGAGGCGCTTCCCTGAAACCCCGGCTGAGAGACCGCGAGGCTGCAAGCAGATC 180
QY 181 ATGAAGCACTGCAAGCCCGCTGCAAGCCGCAAGAGAGCTGCGAGCTGTACAAC 240
DB 181 ATGAAGCACTGCAAGCCCGCTGCAAGCCGCAAGAGAGCTGCGAGCTGTACAAC 240
QY 1072 CGCGTGTGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1128
DB 181 ATACAAAGCTACCAAGCTCTCCAGACAGAGAGAGAGAGAGAGAGAGAGAGAG 240
```





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 1, 2004, 10:43:01 ; Search time 3669 Seconds  
(without alignments)  
12281.829 Million cell updates/sec

Title: US-09-475-704A-4  
Perfect score: 1509  
Sequence: 1 atggcgccgcgcgcagcat.....gcggccctgagccagtaa 1509

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues  
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inh:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	83.2	5.5	933	28	CC391778 PUH1J70TB
2	81.6	5.4	1219	14	CD500131 CDA43-C08
3	81	5.4	749	13	BQ744212 WHB4113_A
4	80.6	5.3	1674	10	BB636712 rockefell

C	5	80.2	5.3	746	14	CK215205	CK215205 FGAS02715
C	6	79.8	5.3	1201	13	BX356664	BX356664 BX356664
C	7	79.8	5.3	2299	11	AY106831	AY106831 Zea mays
C	8	78.4	5.2	1126	14	CK211108	CK211108 FGAS02294
C	9	77.4	5.1	1064	14	CK163121	CK163121 FGAS01573
C	10	77	5.1	1138	29	CG351928	CG351928 OG3AG53TH
C	11	76.8	5.1	1138	14	CK163513	CK163513 FGAS01614
C	12	76.4	5.1	759	10	BF259495	BF259495 HVSMEF001
C	13	76.4	5.1	869	14	CK159167	CK159167 FGAS04056
C	14	76.4	5.1	1308	11	AY104577	AY104577 Zea mays
C	15	75.4	5.0	883	29	CG274385	CG274385 OG2BH19TV
C	16	75	5.0	509	12	BG314500	BG314500 WHE2495_B
C	17	74.4	4.9	1170	14	CD496623	CD496623 CDA23-C11
C	18	74.4	4.9	2598	11	AY103647	AY103647 Zea mays
C	19	74.2	4.9	649	14	CB870888	CB870888 HCL5J04W
C	20	74	4.9	637	12	BI960118	BI960118 HVSMEF002
C	21	73.8	4.9	562	10	BF484304	BF484304 WHE2321_F
C	22	73.2	4.9	636	14	CD932185	CD932185 GR45_117C
C	23	73.2	4.9	694	12	BJ261281	BJ261281 BJ290713
C	24	72.6	4.8	575	13	BQ842001	BQ842001 WHE2987_G
C	25	72.6	4.8	601	12	BJ290713	BJ290713 BJ290713
C	26	72.6	4.8	652	12	BJ292215	BJ292215 BJ292215
C	27	72.6	4.8	683	14	CD905815	CD905815 G468_102P
C	28	72.6	4.8	700	12	BJ295604	BJ295604 BJ295604
C	29	72.6	4.8	712	13	BQ842240	BQ842240 WHE2990_E
C	30	72.6	4.8	718	12	BJ297319	BJ297319 BJ297319
C	31	72.6	4.8	753	12	BJ298801	BJ298801 BJ298801
C	32	72.6	4.8	1150	14	CK161908	CK161908 FGAS01448
C	33	72.6	4.8	1197	14	CK161927	CK161927 FGAS01450
C	34	72.2	4.8	718	14	CD938268	CD938268 OV_109I23
C	35	72.2	4.8	718	14	CD938289	CD938289 OV_109K04
C	36	72	4.8	708	14	CD878344	CD878344 AZO4_102I
C	37	72	4.8	731	13	BQ752847	BQ752847 WHE4119_G
C	38	71.8	4.8	1491	11	AY104431	AY104431 Zea mays
C	39	71.6	4.7	655	12	BJ300074	BJ300074 BJ300074
C	40	70.8	4.7	1327	10	BB636532	BB636532 rockefell
C	41	70.6	4.7	671	12	BI718023	BI718023 1031023D0
C	42	70.6	4.7	738	13	BU295509	BU295509 603608565
C	43	70.4	4.7	764	13	BQ804819	BQ804819 WHE3559_C
C	44	70.2	4.6	648	10	BE517305	BE517305 WHE0616_A
C	45	69.6	4.6	690	14	CD495520	CD495520 CDA17-A10

ALIGNMENTS

RESULT 1  
CC391778 933 bp DNA linear GSS 19-MAY-2003  
LOCUS PUH1J70TB\_ZM\_0.6\_1.0 KB Zea mays genomic clone ZMBMT458K20,  
DEFINITION genomic survey sequence.  
ACCESSION CC391778  
VERSION CC391778.1 GI:30871868  
KEYWORDS GSS.  
SOURCE Zea mays  
ORGANISM Zea mays  
REFERENCE Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,  
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and  
Bennetzen,J.  
TITLE Maize Genomics Consortium  
JOURNAL Unpublished (2003)  
COMMENT Other\_GSSs: PUH1J70TD  
Contact: Cathy Whitelaw  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whitelaw@tigr.org  
Seq primer: TR

Class: sheared ends.  
 FEATURES location/Qualifiers  
 source 1..933  
 /organism="Zea mays"  
 /mol\_type="genomic DNA"  
 /strain="B73"  
 /db\_xref="taxon:4577"  
 /clone="ZM087A458K20"  
 /clone\_lib="ZM 0.6 1.0 KB"  
 /note="Vector: PCR4-TOPO; Site\_1: EcoRI; 0.6-1.0 kb high  
 Cot selected genomic DNA library"

## ORIGIN

Query Match 5.5%; Score 83.2; DB 28; Length 933;  
 Best Local Similarity 45.7%; Pred. No. 1;  
 Matches 329; Conservative 0; Mismatches 388; Indels 3; Gaps 1;

```

QY 189 GCTGAGCCCGCCCTGACAGCCGGCACCAGAGAGCTGCGACGCTGTACAACACCGTGGC 248
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 17 GCCGAGGGCGCGGTGTGCGGCACCGAAGAGCGCCGTGAGCAAGCAGCGCTGTGT 76

QY 249 CACCTGTACTGCTGTCACCGCGGCATGAGTCCGCGACACCAAGAGGCGCTGGACAA 308
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 77 CAGGTGCACAGCACCAGAGAGTTCAGACGCGCTCAAGGCCCAAGAACCGGCTGTGT 136

QY 309 GATCGAGAGAGCAGAACAGTTCAGCAGAAAGACCCAGCAGGCCAAGAGCGCGACGG 368
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 137 GGTGTGAGTTCGCGCGAGCAGCAGAGAGCAGCAGCAGCAGCAGCAGCAGCAGTGT 196

QY 369 CAAGGTGAGCAGAACTACCCCATCTGTCAGAACTCTGACGGGCCAAGTGTGACCAAGG 428
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 197 GCAGCTCAGCGCAGCAGCTGCGCGCAGCAGTGTGACTTCTGTGCTGTGCGCGACGAGT 256

QY 429 CATCAGCCCCCGCACCCTGAACCGCTGGTGAAGTGTATCGAGAGAGAGCGCTTCAGCCC 488
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 257 GAGAGCCACCGAGAGAGCTGTTCGCGCGGAGAGGCGCTCAAGAGGTGCCCACTTCAACTT 316

QY 489 CGAGGTGATCCCCCATGTTCACCGCCCTGAGCGAGAGGCGGCCACCCCGCAGAGACT--GAA 545
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 317 CTACAAAGGGCCCGAAGGTGACAGAGAGAGAGGCCCATCGGCCCGGAGCGGCTGCGGG 376

QY 546 CACGATGTTGAACACCGTGGCGCGCCACAGAGCGCCCATGACATGCTGAAGACACCAT 605
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 377 CGAGTCTCTACTACGCGCAGCAGCAGCTGCGCGGTGTGTCAGCTGCACTCGCGGAGGA 436

QY 606 CAACGAGAGGCGCGCGAGTGGAGCCGCTGCAACCCCGTGCAGGCGCGCGCGCGCGCCC 665
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 437 CGTGAAGGCGCTCATGACAGCAGCAGCGCGCAGCAAGGGGCAAGCTGTGTGCTGAGAGT 496

QY 666 CGGCCAGATGCGGAGACCCCGCGCGCAGCAGCATGCGCGCGCCACAGCAGCCCTGACAGA 725
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 497 GGGCTCAAGCACTGCGGGCGCTGCGTCAAGGTGTACCCACCGTGTGAAGCTGTGCGG 556

QY 726 GCAGATGCGCTGATGATGACAGCAACCCCGCGTGCCTGAGCGACATCTTACAAGCGGTG 785
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 557 GTCCATGTGTGACAAACACCGTCTTTCGCGCGCATGAAACGGCGACAGAACGACAGCTGCAT 616

QY 786 GATCATCTTGGGCGCTGAACAAGATGTCGCGATGTACAGCCCGTGAAGCATCTTGAACAT 845
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 617 GAGATTCTTCAGGGCCATGAAGATGTGAGGTGCCCACTTTCCTTCATCAGGAGCGG 676

QY 846 CCGCCAGAGGCGCGCAAGAGGCGCTTCGCGAGCTACGTGACCGGCTTCTTCAAGACCTGCG 905
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 677 CCAGATGTGCGCGCGCTACGTGCGCTCGCGCAAGGGGGAGCTGTGGTGAGATCTTCCG 736
  
```

RESULT 2  
 LOCUS CD500131/c 1219 bp mRNA linear EST 12-JUN-2003  
 DEFINITION CDA43-C08.y1d-s SHGC-CDA Gasterosteus aculeatus cDNA clone  
 CDA43-C08 3', mRNA sequence.  
 ACCESSION CD500131  
 VERSION CD500131.1 GI:31427162

KEYWORDS EST.  
 SOURCE Gasterosteus aculeatus (three spined stickleback)  
 ORGANISM Gasterosteus aculeatus  
 Bacteria; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 Acanthomorpha; Acanthopterygii; Percomorpha; Gasterosteiformes;  
 Gasterosteidae; Gasterosteus.  
 1 (bases 1 to 1219)  
 REFERENCES  
 Kingsley, D.M., Peichel, C., Balabhadra, S., Grimwood, J., Dickson, M.,  
 Schmutz, J. and Myers, R.M.  
 Expressed sequence tags from Gasterosteus aculeatus  
 Unpublished (2003)  
 CONTACT: Kingsley, DM  
 HMI and Department of Developmental Biology  
 Stanford University School of Medicine  
 Beckman Center B300, 279 Campus Drive, Stanford, CA 94305-5329, USA  
 Tel: 650 725 5954  
 Fax: 650 725 7739  
 Email: kingsley@cimgm.stanford.edu  
 Plate: 43

## COMMENT

High quality sequence start: 4  
 High quality sequence stop: 792.

## FEATURES

source

1..1219  
 /organism="Gasterosteus aculeatus"  
 /mol\_type="mRNA"  
 /strain="Salinas river, CA"  
 /db\_xref="taxon:69293"  
 /clone="CDA43-C08"  
 /sex="mixed male and female"  
 /tissue\_type="heads and internal organs combined"  
 /dev\_stage="adult"  
 /clone\_lib="SHGC-CDA"  
 /note="Vector: lambda ZAP Express/pBK-CMV; Site 1: EcoRI  
 (5' adaptor); Site 2: XhoI (3' linker primer); The mixed  
 organ cDNA library was generated using the ZAP-CDNA method  
 by Stratagene. First strand cDNA synthesis was primed with  
 a 50 bp linker primer containing an oligo dt sequence  
 preceded by a synthetic XhoI site. 5 prime adaptors were  
 used containing an EcoRI cohesive end. The finished cDNAs  
 were inserted in to the ZAP express vector  
 unidirectionally in the sense orientation with respect to  
 the lacZ promoter of pBK-CMV. An amplified library was  
 prepared from approximately 3 million primary clones in  
 the lambda ZAP Express vector. In vivo excision was then  
 used to generate individual pBK-CMV phagemid clones for  
 EST sequencing."

## ORIGIN

Query Match 5.4%; Score 81.6; DB 14; Length 1219;  
 Best Local Similarity 47.7%; Pred. No. 1.7;  
 Matches 373; Conservative 0; Mismatches 394; Indels 15; Gaps 4;

```

QY 618 CGCCGAGTGGAGCCGCTGCAACCCCGTGCAGGCGCGCCCGTGGCCCCCGGCAATGCG 677
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 917 CGGCTCGTGGGCACTACGGCGCGGGTACGGCGACGCCCTGCTCAACGGCTCGCTGGG 858

QY 678 CGACCCCGCGCGAGCGAC-----ATGCGCGGCGCCACAGCAACCTGCAGAGCAGAT 731
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 857 CGACTACGGCGCGCGAGCGCCCTGCTCAACGGCTGCTGGGCACTACGGCTACGGCGGA 798

QY 732 CGCCTGATGACCAAGCAACCCCGCGTGCCTGCGGCGGACATCTTCAAGCGGTGATCAT 791
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 797 CGCCTGCTCAACGGCTGCTGGGGAATAAGGCTACGGCGCGCTGCTCAACGGCTC 738

QY 792 CCTGGCGCTGAACAAGATGCTGGGATGTACAGCCCGTGAAGATCTTGACATCCGCCA 851
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 737 GCTGGGCGACTACGGCGGACCGCCCTGCTCAACGGCAAGCTGGGCACTACGGCGGCGCGGA 678

QY 852 GGGC---CCCAAGAGCCCTTCCGCGACTACGTGACCGGCTTCTTCAAGACCTGCG---G 905
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 677 CGCCTGCTCAACGGCTGCTGGGGAATAAGGCGCGGCTACGGGAGACGCCCTGCTCAA 618
  
```

[illegible]

RESULT	3
LOCUS	BQ744212
DEFINITION	749 bp mRNA linear EST 17-JUL-2002 WH8A113_A01 A01ZS wheat salt-stressed root cDNA library Triticum aestivum CDNA clone WH8A113_A01_A01, mRNA sequence.
ACCESSION	BQ744212
VERSION	BQ744212.1 GI:21890999
KEYWORDS	EST.
SOURCE	Triticum aestivum (bread wheat)
ORGANISM	Triticum aestivum Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae; Triticeae; Triticum. 1 (bases 1 to 749) Anderson, O.D., Akhunov, E., Chao, S., Crossman, C., Deal, K., Dvorak, J., Iazo, G.R., Pham, J., Rausch, C.J., Wilson, C. and Woo, J. The structure and function of the expressed portion of the wheat genomes - Salt-stressed root cDNA library Unpublished (2002) Contact: Olin Anderson US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center 800 Buchanan Street, Albany, CA 94710, USA Tel: 5105595773 Fax: 5105595818 Email: oanderson@pw.usda.gov Sequences have been trimmed to remove vector sequence and low quality sequence with phred score less than 20 Seq primer: SK primer.
JOURNAL	
COMMENT	
TITLE	
AUTHORS	
REFERENCE	

**FEATURES**  
**SOURCE**

```
1. 749
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="WHE4113_A01_A01"
```

/tissue\_type="Roots"  
 /dev\_stage="Full tillering"  
 /lab\_host="E. coli SOLR"  
 /clone\_lib="Wheat salt-stressed root cDNA library"  
 /note="vector: Lambda Uni-ZAP XR, excised phagemid  
 Bluescript SK(-); Site\_1: EcoRI; Site\_2: XhoI; Hydroponic  
 plants grown to full tillering stage were treated with 150  
 mM NaCl for either 12 hours or 7 days. Root tissues of the  
 plants subjected to both types of treatment were collected  
 separately at University of California, Davis (E. Akhunov  
 and K. Deal in J. Dvorak's Lab). Total RNA was prepared  
 separately from the two samples (12h and 7day treatments),  
 and equal amount of RNA was then pooled. PolyA RNA was  
 purified from the pooled RNA, a cDNA library was made, and  
 the cDNA clones were in vivo excised to give Bluescript  
 SK(-) phagemids in J. Dvorak's lab (E. Akhunov, J. Dvorak)  
 at the University of California, Davis. Colony plating,  
 plasmid DNA preparations and DNA sequencing were performed  
 in the OD Anderson lab (all other authors)."

Query Match	5.4%	Score 81	DB 13	Length 749
Best Local Similarity	48.4%	Pred. NO. 1.8		
Matches 293; Conservative	0	Mismatches 300	Indels 12	Gaps 2

QY	265	CCGGCATCGAGSTCCGCGACACCAAGAGGCTCTGGACAAGATCGAGAGCAGAACACA	328
Db	8	CCTTCAGCCTCGTGCACGCCCAAGCCGCCCTTCAAGCACCCGCGCTTCGGGCCCAAGTGGC	67
QY	329	AGTCCAGCAGAAAGACCCTCCAGGCGCCCAAGAGGCCGACCGCAAGTGAAGCCAACTACC	388
Db	68	GCTTCAACCAAGCGGCCCCAGCTCCCCCAGGCGCGGACGAGAGGTGAGGCGCAAGCGCC	127
QY	389	CCATCGTGCAGAACTGCAGGGCCAGATGTGTGACACAGGCCATTCAGCCCCCGCACCTGA	448
Db	128	GCGAGGCCGAGAAAGAGCGCGCGCCGCGCGAGCGCCACTACAGAACCAACCGCTCCACCC	187
QY	449	ACGCGTGGG---TGAAGTGATCGAGAGAGGCTTCAGCCCCGAGGTGATCCCATGT	505
Db	188	ACCACCAAGGCTTCGCGGCAACCAAGACAACCGCAAGCCCTCCGTGACATCCAGC	247
QY	506	TCACCGCCCTGAGCGAGGCGGCCACCCCGCAGAGCCTTAAACAGATGTGAACAACCGTGG	565
Db	248	CCGACTGACCATCTCTGAGCAGATCCCTTCGCCAACTTCACCAAGCTTCCTTCGCGG	307
QY	566	GCGGCCACCAAGCGCGCCATGACAGATGTGAAGACACCATCAACGAGGAGCGCGGAGT	625
Db	308	TCGCCGACCAAGCCCGAGGAACTGTGCTGTGCGGCGCGCTGCACTGTACGACCGCGCT	367
QY	626	GGGACCGCCTGCACCCCGTGCAGGCGCGCGCCCGTGGCCCCCGGCCAGATGCGGACCCCC	685
Db	368	ACGACCGCGTCAACCCCAAGACGGCCCGCGCCTCGAGCGCTTCAAGAACCGCCAGTTCT	427
QY	686	GCGGCAGGCACATCGCCGCGGCCACCAAGCACCTGC-----AGAGCAGATCGCCT	736
Db	428	TCAGATCAACCAACCAACCGAGACCCCATCATTCGCGCGCTTCGCGAGGAGGACAAGGCCA	487
QY	737	GGATGACCAGCAACCCCCCGTGCCTCGTGGGCGACATCTACAGCGGTGATCATCTGG	796
Db	488	CCGCTTTCGCCACCGACGCCATCTCGCGCGCCTCATGTGACGCCCCCGCAGCATCTCT	547
QY	797	GCCTGAACAGATCGTGCAGATGTACAGCCCCGTGAGCATCTGACATCCGCCAAGGCC	856
Db	548	CCTGGGACATTTGTGTGACAGCGCGCTCGGCAACAAGCTCTTCTTGACAAGCGCGAGGGCT	607
QY	857	CCAG 861	
Db	608	CCGAG 612	

RESULT 4				
BE636712				
LOCUS	BE636712	1674 bp	mRNA	linear EST 03-JAN-2002

DEFINITION rockefeller.0.42 Masticamoeba balamuthi lambda ZAP II library  
Masticamoeba balamuthi cDNA similar to vacuolar ATP synthase  
subunit B (EC 3.6.1.3), mRNA sequence.  
BB636712  
VERSION BB636712.1 GI:9919823  
KEYWORDS EST.  
SOURCE Masticamoeba balamuthi  
ORGANISM Masticamoeba balamuthi  
Eukaryota; Pelobiontida; Masticamebidae; Masticamoeba.  
1 (bases 1 to 1674)  
REFERENCES Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W.,  
Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and  
Phillippe,H.  
The analysis of 100 genes supports the grouping of three highly  
divergent amoebae: Dictyostelium, Entamoeba, and Masticamoeba  
Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)  
21819461  
PUBMED 11830664  
COMMENT Contact: Muller Miklos  
Laboratory of Biochemical Parasitology  
The Rockefeller University  
1230 York Avenue, New York, NY 10021, USA  
Email: mmuller@rockvax.rockefeller.edu  
Insert Length: 1674 Std Error: 0.00  
POLYA=Yes.  
FEATURES  
source Location/Qualifiers  
1..1674  
/organism="Masticamoeba balamuthi"  
/mol\_type="mRNA"  
/strain="ATCC 30984"  
/db\_xref="taxon:108607"  
/clone\_lib="Masticamoeba balamuthi lambda ZAP II library"  
/note="syn: Phreatamoeba balamuthi"  
ORIGIN  
Query Match 5.3%; Score 80.6; DB 10; length 1674;  
Best Local Similarity 44.1%; Pred. No. 2.4;  
Matches 615; Conservative 0; Mismatches 756; Indels 22; Gaps 6;  
QY 59 GCCTGCGCCCGCGGCAAGACACTACATGCTGAAGCACTGTGTGGCCAGCCGCG 118  
DB 136 GCCCTCTGCTCTCTCTCGACACATCAAGAGCCGCGAGTTCCGAGAGATCGTCACTGA 195  
QY 119 AGCTGAGAGGCTTGGCCCTGAACCCCGGCTGTGAGAGCCGCGAGGCTGCAAGCAGA 178  
DB 196 CGCTGGGCGACGGGAGCGTCCCGCGGAGGCTGTGAGAGTGGCCGCGAACAAGGCGC 255  
QY 179 TCATGAAGCACTGCAAGCCCGCTTGCAGACCGGCAACGAGAGCTGCGCAGCTGTACA 238  
DB 256 TCGTGCAGGCTTTGAGGGGCACTGCGGCAATGACGCCAAGAACACGCACTGCGAGTTCA 315  
QY 239 ACACCGTGGCCACCTGTACTGCTGTGCAAGCGCGGCATGAGGTCCGCGACACCAAGA-- 296  
DB 316 CGGGCGACATCATGCGCATCCCGCTCAGCGAGGACATGCTGGCGCGCTTCAAGGCT 375  
QY 297 -GGCCCTGACAGATCGAGAGAGAGACAGAACAGTCCCAAGCAAGACCCAGAGGCCA 355  
DB 376 CGGGCGCGCGCTGCAAGAGGCGCGCGGTCTATGGCGGAGACTTCTCTGACATCCAGG 435  
QY 356 AGAAGGCCGACGGCAAGGTGAGCCAGAACTACCCCATCTGTGCAAGACTGCAAGGCCAGA 415  
DB 436 GCCAGCCCATCAACCCGTTACAGCGTGTATCCCGAGAGATGATCCAGACGGGCATCA 495  
QY 416 TGGTGACCAAGGCGCATCAGCCCGCGCAACCTGAAAGCTGGGTGAAGGTGATGAGAGA 475  
DB 496 GCGCCATCGACAGATGACTGATCGCCCGTGGCCAGAAAGTCCCGCTGTTCAAGCGCG 555  
QY 476 AGGCTTCAAGCCCGGAGGTATCCCATGTTCAACCGCCCTGAGAGAGGCGGCCACCCCG 535  
DB 556 CCGGTCTGCGGCAACAGAGATTGCGCGCAGATCTGCAAGGACGCGCGCTGTGCGCC 615  
QY 536 AGGACCTGAACAGATGTTGAACACCGTGGGCGGCCACAGGCGGCCATGAGATGCTGA 595

DB 616 TCCAGAGCAAGAAGTTCAGAGACGACCAAGAGCAACTTCGCATCTGTTCGCG 675  
QY 596 AGGACACCATCAACGAGAGGCGCGGAGTGGAGCCGCTGCAACCCGTCAGCGCGCC 655  
DB 676 CCATGGGTCAACATGAGACTGCGAGGTTCTTCGCGAGAGCTTCAGAGAGAACGGCT 735  
QY 656 CCGTGGCCCGCGGCAAGTGGCGGACCCCGCGGAGAGCATGCGCGGCCACAGCA 715  
DB 736 CGATGAGCGGTTCACGCTGTTCTCAACCTCGCAACCAACCCACGATCGAGCGCATCA 795  
QY 716 CCCTGACAGAGCATGCGCTGATGACCAAGCAACCCCGGTGCGGTGGGACATCT 775  
DB 796 TCACGCGCGCTG--GCGCTCAGACGCGCGAGTACCTGCGGTACAGTGGAGATGC 852  
QY 776 ACAAGCGGTGATCATCTGGGCT-----GAAACAAGTGTGGGATGTACAGCCCG 829  
DB 853 ACGTGTGTGATCTGACGACATATGCTGCTGTAACGCCGATGCCCTGCTGAGTCTCG 912  
QY 830 TGAGCATCTGACATCCGCGCAGGCGCCCAAGAGCCCTTCGCGGACTACGTGACCGCT 889  
DB 913 CGGCGCGTGAAGAGTGCCTGGGCGTGTGTTACCCGCGTACATGTACAGAGATCTCG 972  
QY 890 TCTTCAAGACCTGCGCGCGGAGGAGCCAGGAGAGAGAGTGAAGAACTGATGACGAGA 949  
DB 973 CGACCATCTACAGCGGTGCGCGCGTGTG--TGAAGGAGAGAGAGCGCTGATCAGCAGC 1029  
QY 950 CCCTGTGTGAGAGAGCGCAACCCCGACTGCAAGACCATCTGCGCGCTCTGCGCCCG 1009  
DB 1030 TCCCATCTCTACGATGCGCCCAAGAGACATCAACCCCATCTCTGACCTCAGCGGCT 1089  
QY 1010 GCGCCACCTGAGAGAGATGATGACCGCTGCGGAGGCGGTGGCGCGCCCGCAAGG 1069  
DB 1090 ACATCACCGAGGCGGACGATCTACGTGACAGGCACTGCAACAAGAGATCTACCGC 1149  
QY 1070 C---CGCGTGTGCGCGGAGGCGATGAGCCGCAACAGCGTGAACATCATGATGAGA 1126  
DB 1150 CCATCAACGTGTGCGCGCTGCTGCTGAGCTGATGAAGAGCGGCATCGCGAGGAGATGA 1209  
QY 1127 AGAGCAACTTCAAGGCGCGCGCGCAACGTCAAGTCTTCAACTGCGGCAAGAGGCGC 1186  
DB 1210 CGCGGACGACCACTCGAGAGTGTGAACCACTGTACGCGCACTACGCCATGCGCAAG 1269  
QY 1187 ACATGCGCAAGAACTGCGCGCGCGCGCGCAAGAGGCTGTGAAGTGCAGAGAGG 1246  
DB 1270 ACACGATGGCCATGAAGGCTGTGTGCTGAGAGAGGCGCTGTGCGCGAGGAGAGCTCG 1329  
QY 1247 GCCACCAAGTGAAGACTGCAACCGAGCGGCACTTCTTGGGCAAGATCTGGCCA 1306  
DB 1330 CCTTGAAGTTCCTTCGAGAGATTGAGCAAGAGTTTCATCAGCGAGGCTACACAGAGTCC 1389  
QY 1307 GCCACAAGGCGCGCGCGCACTTCTGCAAGAACCGGAGCGCGCGCGCCCAACCG 1366  
DB 1390 GGCACATCTACAGTGTGCTGACATGCTGAGAGCTCTGCG---CAAGTTCGCCAAG 1445  
QY 1367 TGCCACCGCGCGCGCGCGAGAGCTTCCGCTTGGAGAGAGCAACCCCGCCCAAGC 1426  
DB 1446 AGCTCTCAGCGGTATCAGGAGAGAGACCTTCAACAAGTACTACAGAGGAGCGCAAGN 1505  
QY 1427 AGGAGCCCAAGGA 1439  
DB 1506 AGAGACCAAGTA 1518  
RESULT 5  
CK215205/c 746 bp mRNA linear EST 09-DEC-2003  
LOCUS FGAS027158 Triticum aestivum FGAS: Library 6 CAP GATE 1 Triticum  
DEFINITION aestivum cDNA, mRNA sequence.  
ACCESSION CK215205  
VERSION CK215205.1 GI:39621309  
KEYWORDS EST.  
ORGANISM Triticum aestivum (bread wheat)  
Triticum aestivum





```

QY 954 GCTGTGCAAGACGCCAACCCTGCAAGACCATCTGCGCCCTCTCGCCCGCGCC 1013
Db 1141 SVVSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 1082
QY 1014 CACCTTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1073
Db 1081 ASVSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 1022
QY 1074 CGTGTGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1133
Db 1021 SSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 962
QY 1134 CTTCAAGGCGCCCGCGCAAGTCAAGTCTTCAAGTCTTCAAGTCTTCAAGTCT 1193
Db 961 MSSAVSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 902
QY 1194 CAGAACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1253
Db 901 SAAAVASAAVAAVAAVAAVAAVAAVAAVAAVAAVAAVAAVAAVAAVAAVAA 842
QY 1254 GATGAGGAGTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1313
Db 841 SSVAVSAAVAAVAAVAAVAAVAAVAAVAAVAAVAAVAAVAAVAAVAAVAA 782
QY 1314 GGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1373
Db 781 VSSVSVSSSVAAVAAVAAVAAVAAVAAVAAVAAVAAVAAVAAVAAVAAV 722
QY 1374 CGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1433
Db 721 SAAVAAVAAVAAVAAVAAVAAVAAVAAVAAVAAVAAVAAVAAVAAVAA 662
QY 1434 CAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1493
Db 661 SASSSAAVAAVAAVAAVAAVAAVAAVAAVAAVAAVAAVAAVAAVAAVAA 602
QY 1494 CCCCCCTGAGCCAGTAA 1509
Db 601 NSSNSNANNNNANAAA 586

```

```

RESULT 7
AY106831 2299 bp mRNA linear HTC 16-OCT-2002
DEFINITION Zea mays PC0142079 mRNA sequence.
ACCESSION AY106831
VERSION AY106831.1 GI:21209909
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 2299)
Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S.,
Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
Unpublished (2002)
2 (bases 1 to 2299)
Coe, E.H.

```

```

JOURNAL Direct Submission
REFERENCE Submitted (25-APR-2002) Maize Mapping Project, University of
AUTHORS Missouri, Columbia, MO 65211, USA
TITL If you are interested in getting corresponding physical clones,
JOURNAL these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
sequence is either Virginia Walbot, Stanford or Pat
Schubert, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.
COMMENT

```

```

FEATURES
SOURCE Location/Qualifiers
1..2299

```

```

/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:638374"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"

```

## ORIGIN

```

Query Match 5.3%; Score 79.8; DB 11; Length 2299;
Best Local Similarity 43.9%; Pred. No. 3.1;
Matches 495; Conservative 0; Mismatches 617; Indels 15; Gaps 3;

```

```

QY 122 TGAAGGCTTGGCCCTGAAACCCCGGCTGCTGAGAGACCGCGAGGCTGCAAGCATCA 181
Db 767 TCGAGCGCGCGCGCATATGAGACACATCTGATGAGCTCTTATGAGACGCGCA 826
QY 182 TGAAGAGCTGAGCCCGCGCTGAGAGACCGCGAGAGCTGCGCAGCTGTACACA 241
Db 827 AGAGAGTGAAGCCCATGAGACCGCGCTGCTGAGAGAGAGAGAGAGAGAGAGAG 886
QY 242 CCGTGGCCACCTGTACTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 301
Db 887 CCGTGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 946
QY 302 TGAAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 361
Db 947 AGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1006
QY 362 CCGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 421
Db 1007 TGAAGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1066
QY 422 ACCAGGCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 481
Db 1067 CCGTGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1126
QY 482 TGAAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 541
Db 1127 ACAACAGCGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1186
QY 542 TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 601
Db 1187 TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1246
QY 602 CCATCA-----ACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 655
Db 1247 CCATCAACCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1306
QY 656 CCGTGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 715
Db 1307 TCGTCTCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1366
QY 716 CCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 775
Db 1367 ACAAGTGTGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1426
QY 776 ACAAGCGTGAATCATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 835
Db 1427 CCGTCAAGAGCTGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1486
QY 836 TCGTGAATCATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 895
Db 1487 ACTTCCACAGCGCGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1546
QY 896 AGACCTTGGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 949
Db 1547 TGTACGGGTACTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1606

```





QY 759 GCCCGTGGGCGGACATCTACAGCGGTGATCATCTGGGCTGMAACAGATCTGGCGAT 818  
DB 596 GCGGTGGAGCCCGCGGTGTGTGGCGCGCGCGCGCGCGAGCAGAGGCCAAGAGCGC 537  
QY 819 GTACAGCCCCGTGAGCATCTTGACATCCGCCAGGGCCCCAAGAGAGCCCTTCCGCGACTA 878  
DB 536 CAGCAGCTCCCGGTCCAGAGACTTTCGGGCGCGCGCGCGCAATGGTGGGGCCAGAGAGT 477  
QY 879 CGTGAGCCGCTTCTTCAAGACCTTCGCGCGCGCGAGCAGGCCACCCAGAGAGTGAAGACTG 938  
DB 476 GAGAGAGAGATGCGCGCGCTGTGCGCGGTGCTGCGGCGCA--AGAGCAGCAGAGATA 420  
QY 939 GATGACCGGACCTGCTGTGTGAGAGAGCCCAACCCGACTGCAAGACCATCTGCGCGC 998  
DB 419 CTTGAGGGGTGGGAGACTGTGTGCGCGGTCAACCGCG--GCTTGGCGGTGGCGGGGCC 363  
QY 999 TCTCGGCCCCGCGCGCACCCCTGAGAGAGATGATGACCGCTGCGAGGGCGTGGCGGCC 1058  
DB 362 CCGCTTGGCCCGCGCGCGCGCGGTGCGCAGGCTTTCATCGGCGGAGCGCGCTGGGC 303  
QY 1059 CCGCCACAAGGCCCGCGGTGTGCGCGAGCGATGAGCGCAGGCCAAGCGTGAACATCAT 1118  
DB 302 GTTCGGCGCGCGGTGACGTGCGCGCGCTGGCGCGCGGTCAACAGGTGAGCAGCG 243  
QY 1119 GATGACAGAGAGCAACTTCAAGGGCCCCCGCGCAAGTCAAGTCTTCAACTGCGGCA 1178  
DB 242 CCGCAGACTGGGCACTGTCTGTTCAGAGTGTGCGCAACTGCGCGCTTCTCAACGCAACT 183  
QY 1179 GAGGGCCACATCGCCCAAGACTGCGCGCGCCCCCGCAAGAGGGCTGTGAAGTGGG 1238  
DB 182 GAGGAGAGACATCGAGCGCGCGCTGGCGAGCCCGAGGTGAGCGCGGAGAGACGGGA 123  
QY 1239 CAAGAGGGGCCACCATGAGAGACTGCAACGAGCGCAGGCCCAACTTCTGGGCAAGT 1298  
DB 122 CGTGTTCGGAGCAGAGGTGGCGCTGTGTGCGCGCGAGCGGTGAGACTCCGCGACTT 63  
QY 1299 CTGGC 1303  
DB 62 CAGGC 58

RESULT 11  
CK163513 1138 bp mRNA linear EST 05-DEC-2003  
LOCUS FGAS016142 Triticum aestivum FGAS: Library 4 Gate 8 Triticum  
DEFINITION aestivum cDNA, mRNA sequence.  
ACCESSION CK163513  
VERSION CK163513.1 GI:36993817  
KEYWORDS EST.  
SOURCE Triticum aestivum (bread wheat)  
ORGANISM Triticum aestivum  
REFERENCE Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
1 (bases 1 to 1138)  
Allard, F., Crosby, W.L., Danyluk, J., Budes, F., Frick, M., Gaudet, D.,  
Geneswein, B., Graf, R., Gulick, P., Hrycan, L.D., Laroche, A.,  
Links, M.G., McCarthy, E.L., Monroy, A., Muzak, I., Nilson, D.,  
Penniket, C., Roach, J.L. and Sarhan, P.  
Functional Genomics of Abiotic Stress in Wheat and Canola Crops  
Unpublished (2003)  
CONTACT: Wm L Crosby  
Bioinformatics  
University of Saskatchewan, Department of Computer Science  
1C101 Engineering Building, 57 Campus Drive, Saskatoon,  
Saskatchewan, S7N 5A9, Canada  
Tel: 306 966 1769  
Fax: 306 966 2033  
Email: fgas\_estecs.usask.ca  
This sequence is the direct result of the Base calling software  
phred (default parameters). It is the raw base calls. To aid in the  
identification of the high quality insert the software Lucy

FEATURES  
source

(default parameters) has been run on this sequence. Lucy identified  
the region (8,827).  
Plate: LAB008 row: F column: 04.  
location/Qualifiers  
1..1138  
/organism="Triticum aestivum"  
/mol\_type="mRNA"  
/db\_xref="taxon:4565"  
/note="Organ: Crown and leaf; Vector: pCMV.SPORT6; Aerial  
parts (crown and leaf) of wheat cultivar Norstar during  
dehydration stress. 8 mRNA populations were combined  
before constructing the library. The first four come from  
removing plants from vermiculite (7 day old plants) and  
incubating them at 20C on the bench without water for 1,  
2, 3 and 4 days. The last four come from plants grown in  
soil in a growth chamber after watering is terminated.  
Four samplings were taken in a two week period; the first  
after wilting was observed and the last, two weeks later,  
consisted of live crown and stem tissue (leaf tissue was  
yellow and dead). First strand synthesis in this library  
was done in the presence of methylated dCTP thereby  
protecting from internal cleavage with NotI."

ORIGIN

Query Match 5.1%; Score 76.8; DB 14; Length 1138;  
Best Local Similarity 47.9%; Pred. No. 6.3;  
Matches 283; Conservative 0; Mismatches 303; Indels 5; Gaps 2;

QY 124 GAGGCTTGGCCCTGAACCCCGGCTGTGAGACCGCGAGGGCTGCAAGCAGATCATG 183  
DB 232 GAGATCAACCGCAACCAACCATATGAACTGATGACAGCTGTGGCGAGCGCCAGTTTC 291  
QY 184 AAGCACTGACGCCCGCTGACAGACCGCAACGAGAGCTGCGAGCTGTACAACACC 243  
DB 292 CTCAACATGTCTCTCAAGCTCATCGGCGCAAGAGACCATGAGATCGGCTGTACACC 351  
QY 244 GTGGCCACCTGTACTGCTGTGACAGCGCGCATGAGGTCCGACACCAAGAGGCCCTG 303  
DB 352 GGCTACTTCCTGTGTGCGACCGCGCTGCGCATCCCGACAGCGGACCATCTTGGCCATG 411  
QY 304 GACAAGATCGAGAGAGAGACAGAACAGTCCAGCAGAGACCAGAGGCCAAGAGGCC 363  
DB 412 GACATCAACCGCAGAACTACAGAGTGGGGCTGCGGTGATCGAAGAGCGCGGTGGCG 471  
QY 364 GACGGCAAGGTGAGCCCAACTACCCATGCTGACAGAACTGACAGGGCCAGATGTGAC 423  
DB 472 CACAGATGACTTCCGAGAGGCCCGCGCTGCGCTGTGAGCGCGCTGTGAGAGAC 531  
QY 424 CAGGCCATCAGCCCCCGCACCTGAGAGCGCTGGGTGAAGTGTATGAGAGAGGCCCTTC 483  
DB 532 GAGGCCAACACGAGACTTGTGACTTGTGTGAGCGCCGACAGAGCAACTACTTC 591  
QY 484 AGCCCGAGGTGATCCCATGTTCACCGGCTGAGAGAGGCCACCCCCAGAGACTTG 543  
DB 592 AACTACCAAGAGCGCTCATGAGCTGTCAAGTGGCGGCTCTCGGCTACGAC--- 648  
QY 544 AACACATGTGAACACCGTGGCGCGCACAGCGCCATGACATGTGAAGAGACACC 603  
DB 649 AACACCTCTGAACGCGCTGCTGTGCTCCCGCGAGCGCCCATGCGCAAGTACATC 708  
QY 604 ATCAAGAGAGAGCGCGAGTGGACCGCTGCACCCGTGACAGGCCGCG--CCGTTG 661  
DB 709 CGTACTACCGGAGCTGTCTGTGAGCTCAACAGGCGCTCGCGGACAGCGCGTGG 768  
QY 662 CCCCCGCGAGATGCGGACCGCCCGCGGACGACATCGCGCGCCACCA 712  
DB 769 AGATCTGCAAGTCCCGGTGGCGAGGATACCTCTGCGCGCGGCCA 819

RESULT 12  
BF259495 759 bp mRNA linear EST 22-OCT-2001  
LOCUS BF259495



**DEFINITION** HVSMF0019H19f Hordeum vulgare seedling root EST library HVCNDA0007  
(Et isolated and unstressed) Hordeum vulgare subsp. vulgare cDNA

**ACCESSION** BF259495  
**VERSION** BF259495.2 GI:13120022

**KEYWORDS** EST.

**SOURCE** Hordeum vulgare subsp. vulgare  
Hordeum vulgare subsp. vulgare  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Poideae; Triticeae; Hordeum.

**REFERENCE** 1 (bases 1 to 759)  
Wing, R., Close, T.J., Kleinholts, A., Wise, R., Begum, D., Frisch, D.,  
Yu, Y., Henry, D., Palmer, M., Rambo, T., Simmons, J., Choi, D.W.,  
Penton, R.D., Oates, R. and Main, D.  
Development of a genetically and physically anchored EST resource  
for barley genomics: Morex unstressed seedling root cDNA library  
Unpublished (2001)  
On Nov 16, 2000 this sequence version replaced gi:11188608.

**JOURNAL** Contact: Wing RA  
Clemson University Genomics Institute  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 7288  
Fax: 864 656 4293  
Email: rwing@clemson.edu

**COMMENT** Total hg bases = 388  
Seq primer: AATTACCCCTCACTAAGG  
High quality sequence stop: 758.  
Location/Qualifiers

**FEATURES**

**source**

1..759

/organism="Hordeum vulgare subsp. vulgare"

/mol\_type="mRNA"

/cultivar="Morex"

/sub\_species="vulgare"

/db\_xref="taxon:112509"

/clone="HVSMF0019H19f"

/issue\_type="Seedling root"

/lab\_host="TJC121"

/clone\_lib="Hordeum vulgare seedling root EST library  
HVCNDA0007 (Et isolated and unstressed)"

/note="Vector: lambdaZAP; Site 1: EcoRI; Site 2: XhoI;  
Seeds were surface sterilized then germinated under axenic  
conditions in the dark at room temperature on filter paper  
with water, mycstatin and cefotaxime in covered  
crystallization dishes. Five-day old seedling roots were  
then harvested, total RNA was prepared, poly(A) RNA was  
purified, one primary unamplified cDNA library was made,  
and 1 million plv were in vivo excised to give Bluescript  
SK(-) cDNA phagemids. These steps were performed in the Tu  
Close laboratory at the University of California,  
Riverside (Choi, Close, Penton). Phagemids were plated and  
picked at the Clemson University Genomics Institute (CUGI)  
(Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA  
preparations, DNA sequencing and sequence analysis were  
performed at CUGI (Wing, Yu, Frisch, Henry, Simmons,  
Oates, Rambo, Main). The sequence has been trimmed to  
remove vector sequence and contains a minimum of 100 bases  
of phred value 20 or above. For more details on library  
preparation and sequence analysis see  
http://www.genome.clemson.edu/projects/barley. To order  
this clone see http://www.genome.clemson.edu/orders\_Albo  
see Close TJ, Wing R, Kleinholts A, Wise R (2001)  
Genetically and physically anchored EST resources for  
barley genomics. Barley Genetics Newsletter 31:29-30.  
(http://wheat.pw.usda.gov/g9pages/bgn/31/cover.html)"

**ORIGIN**

Query Match 5.1%; Score 76.4; DB 10; Length 759;  
Best Local Similarity 47.8%; Pred. No. 6.5;  
Matches 283; Conservative 0; Mismatches 303; Indels 6; Gaps 2;

124 GAGGCTTGGCCCTGAACCCCGGCTGTGAGACCGCGAGGCTGCAAGCATCATG 183

DB 13 GAGATCACCGCCCAACCAACCCATGAACTGTATGACGACGTGGGCGAAGAGGGCCAGTTTC 72

QY 184 AAGCAGCTGCAGCCCGCCCTGCGAGACCGGCAGGAGCTGCCAGCTGTATCAACACC 243

DB 73 CTCAACATGTCTGCTCAAGCTCATTCGGCGCCAGAAAGACCATGAGATCGGCTTACACG 132

QY 244 GTGGCCACCTGTATCTGCGTGCAGCCCGGCTGAGGTCCGCCACCAAGAGGCGCTTG 303

DB 133 GGCTACTCCCTGCTCGCCACCGCGCTCGCCATCCCGACGACCGGACCATCTTGGCCATG 192

QY 304 GACCAAGATCGAGAGAGAGACAGAACTCCCAAGAGAGAGACCAAGAGAGAGAGCC 363

DB 193 GACATCAACCGGAGAACTACGAGCTGGGCTGCTGCTGATCGAAGAGCGCGCTGGCG 252

QY 364 GAGGCAAGGTGAGCCAGAACTACCCCATCTGTCAGAACTGACAGGCGCAGATGTGAC 423

DB 253 CACAAGATGACTTCGCGAGGGCGCGCTCCCTGCTGACGCTCTCTGAGAGAC 312

QY 424 GAGGCTATCAGCCCGGCGACCTGTAAGCTGCTGAGTGAAGTGAATCGAGAGAGCTTC 483

DB 313 GAGGCCAACGAGGACCTTGCACTTCGCTGCTGAGCGCGCAAGAGCAACTACCTC 372

QY 484 AGCCCCGAGGTGATCCCATGTTCAACCGCCCTGAGCGAGGCGCCACCCCGACAGCTG 543

DB 373 AACTACCAAGAGGCTCATGAACTCGTCAAGTTCGCGGCTCTCTGCTACAC--- 429

QY 544 AACACGATGTGAACACCGTGGCGGCCACCGAGCGCCATGACATGCTGAAGACACC 603

DB 430 AACACCTCTGGAACGCGCTCCGCTGTGCTCCCGCGAGCGCCCATGCGCAATACATC 489

QY 604 ATCAACGAGAGCGCGCGAGTGGAGACGCTGCAACCCCGTGCAGGCGCGCC---CGTG 660

DB 490 CGCTACTACCGGCACTTGCTGCTGCACTCAACAGAGCCCTGCGCGCGACCGCGCTN 549

QY 661 GCCCGCGGCGAGTGGCGAGCCCGCGGAGGACATCGCGCGGCCACCA 712

DB 550 GAGATCTGCGAGCTCCCGCTGCGGAGCGGACATCACTCTGCGCGCGGCCA 601

**RESULT 13**

CK159167/c 869 bp mRNA linear EST 05-DEC-2003

**LOCUS** FGAS040564 Triticum aestivum FGAS: Talt5 Triticum aestivum cDNA,  
mRNA sequence.

**DEFINITION** CK159167

**ACCESSION** CK159167.1 GI:38985053

**VERSION** EST.

**KEYWORDS** Triticum aestivum (bread wheat)

**SOURCE** Triticum aestivum

**ORGANISM** Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Poideae; Triticeae; Triticum.

**REFERENCE** 1 (bases 1 to 869)  
Allard, F., Crosby, W.L., Danyluk, J., Budes, F., Frick, M., Gaudet, D.,  
Genswein, B., Graf, R., Gulick, P., Hyman, L.D., Laroche, A.,  
Links, M.G., McCarthy, E.L., Monroy, A., Muzak, I., Nilson, D.,  
Peniket, C., Roach, J.L. and Sarhan, F.  
Functional Genomics of Abiotic Stress in Wheat and Canola Crops  
Unpublished (2003)

**AUTHORS** Contact: Wm L Crosby

**TITLE** Bioinformatics

**JOURNAL** University of Saskatchewan, Department of Computer Science  
1C101 Engineering Building, 57 Campus Drive, Saskatoon,  
Saskatchewan, S7N 5A9, Canada  
Tel: 306 966 1769  
Fax: 306 966 2033

**COMMENT** Email: fgas\_est@cs.usask.ca  
This sequence is the direct result of the Base calling software  
phred (default parameters). It is the raw base calls. To aid in the  
identification of the high quality insert the software Lucy  
(default parameters) has been run on this sequence. Lucy identified  
the region [128,636].



Plate: Tal537 row: N column: 23.

## FEATURES

source

1. 869  
/organism="Triticum aestivum"  
/mol\_type="mRNA"  
/cultivar="wheat line PI 178383"  
/db\_xref="taxon:4565"  
/lab\_host="DHS alpha"  
/clone\_lib="Triticum aestivum FGAS: Tal5"  
/note="Organ: Crown; Vector: pGEM-T; SSH (suppression subtractive hybridization) cDNA library from genotype PI178383 cold hardened at 2 C for 21 days and 49 days (equal amount of cDNA pooled together before subtraction, tester) and subtracted against genotype Norstar cold hardened at 2 C for 1 day (24 H) (driver). Modified Smart cDNA (Clontech) priming and non-directional cloning"

## ORIGIN

Query Match 5.1%; Score 76.4; DB 14; Length 869;  
Best Local Similarity 45.2%; Pred. No. 6.7;  
Matches 281; Conservative 0; Mismatches 341; Indels 0; Gaps 0;

QY 663 CCCCCGCGAGATGCGGAGACCCCGCGGAGCATCGCCGCGCCACGACCCCTGCA 722  
DB 780 CCACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 721  
QY 723 GGAGCAGATGCGCTTGATGACCAACCCCGCGGAGCATCTACAAGCG 782  
DB 720 CAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 661  
QY 783 GTGATCATCTGGGCTGGAACAAGATCGTGGGATGTACAGCCCGTGAGCATCTGGA 842  
DB 660 CAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 601  
QY 843 CATCCGCGAGGCGCCCAAGAGCCCTTCCGCGACTACGTGACCGCTTCTCAAGACCT 902  
DB 600 CAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 541  
QY 903 GCGCGCGGAGCAGGCGCACCGGAGCATGGAAGTGAAGTGAAGTGAAGTGAAGTGA 962  
DB 540 CAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 481  
QY 963 GAACGCGCAACCCGACTGCAAGACCATCTGCGGCTCTCGGCGCCGCGCCCTGGA 1022  
DB 480 CAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 421  
QY 1023 GGAGATGATGACCGCTGCGAGGGGCGTGGGCGCGCCGCGGCGCGCTGCGC 1082  
DB 420 CGACGACGACGACGACCAACCAACCAACCAACCAACCAACCAACCAACCA 361  
QY 1083 CGAGGCGATGAGCCAGGCGCAAGCGTGAACATCATGATGAGAAGAGCAACTTCAAGGG 1142  
DB 360 CAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 301  
QY 1143 CCCCCGCGGCAACGTCAAGTGTCTCAACTGCGGCAAGAGGCGCATGCGCAAGAACTG 1202  
DB 300 CAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 241  
QY 1203 CCGCGCGCGCGCGCAAGAGGCTGTGGAAGTGGCGCAAGAGGCGCCACGATGAAGGA 1262  
DB 240 CAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 181  
QY 1263 CTGACCGAGAGCGCCAGGCGCAAC 1284  
DB 180 CAACCAACCAACCAACCGCAAC 159

## RESULT 14

AY104577 1308 bp mRNA linear HTC 16-OCT-2002  
DEFINITION Zea mays PC0137756 mRNA sequence.  
ACCESSION AY104577  
VERSION AY104577.1 GI:21207655

## KEYWORDS

SOURCE

ORGANISM

HTC.  
Zea mays  
Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.

## REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Submitted (25-APR-2002) Maize Mapping Project, University of  
Missouri, Columbia, MO 65211, USA  
If you are interested in getting corresponding physical clones,  
these are publicly available from ZmDB and may be found by BLAST  
searching at MSU, [maizemap.org](http://maizemap.org); ZmDB, [www.zmdb.iastate.edu](http://www.zmdb.iastate.edu); TIGR,  
[www.tigr.org](http://www.tigr.org); or NCBI, [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov). When the source of the  
maize cDNA sequences is either Virginia Walbot, Stanford or Pat  
Schnable, Iowa State, then clones may be requested from ZmDB:  
[www.zmdb.iastate.edu](http://www.zmdb.iastate.edu).

## FEATURES

source

1. 1308  
/organism="Zea mays"  
/mol\_type="mRNA"  
/db\_xref="MaizeDB:638085"  
/db\_xref="taxon:4577"  
/clone\_lib="Maize Mapping Project/Dupont Cornsensus  
Library"  
/note="this sequence is part of a project of EST  
assemblies resulting from the application of public  
contigs to seed Dupont contigs; this resource was  
assembled by Dupont as part of a collaboration for the  
overgo addressing of BACs in conjunction with the Maize  
Mapping Project"

## ORIGIN

Query Match 5.1%; Score 76.4; DB 11; Length 1308;  
Best Local Similarity 44.7%; Pred. No. 7.2;  
Matches 519; Conservative 0; Mismatches 601; Indels 42; Gaps 4;

QY 294 GAGAGCCCTGACAGATCGAGAGAGAGCAGACAAGTCCAGACGAAGACCCAGCAGGC 353  
DB 2 GAGCGGTGGAGAGAGAGAGTGGCGGCTTGAAATCGACGAGCTGATGAGAAAGCCGA 61  
QY 354 CAAGAGGCGCCAGCGCAAGGTGAGCCAGAACTTACCCCATCTGTGACGAACCTGAGGCCA 413  
DB 62 CGGGTTGCGGGGGGTGTTCCCGAGCACAAGTACAGATCGTGAAGCGGCTGAGAACCG 121  
QY 414 GATGTGACACGAGCCATCAGCCCGCCGACCCCTGAAGCGCTGGTGAAGTGAAGGA 473  
DB 122 GAGCACAATCTGCGGATGACCGGGGACCGGCTGAACGACGCTCCGCGCTGAAGAGGC 181  
QY 474 GAGGCTTCAAGCCCGGAGGTGATCCCATGTTTCAACCGCTGAGCGAGGCGCCACCCC 533  
DB 182 GGCATGCGGATGCGGTGAGAGCGGACGCGGCGCGGCGGCGGCGGTGGAATCTGT 241  
QY 534 CCAGACCTGAACACGATGTGAACACCGTGGCGGCGCCACGAGCGCCATGAGATGCT 593  
DB 242 GCTGAGGAGCGCGGCGTGAAGCTGATGTGAGCGCGGTGCTCAACGCGCGCATCTT 301  
QY 594 GAAGGACACCATCAACGAGAGGCGCGGAGTGGGACCGCTTGACACCCCGTGAAGGCGCG 653  
DB 302 CCAGCGCATGAAGAACTACACCATCTACGCGGTGTCATACATCCGATCGTGTGG 361  
QY 654 CCCCCGCGCGCGCGCGAGTGGCGGACCCCGCGGCGGAGCATGCGCGGCGGCGGACGAG 713  
DB 362 CTTCCTGTGCTGCGCTGTGGAAGTTGACTTGGCGCCCTTCATGTGCTCATCAT 421  
QY 714 CACCCTGACGAGAGCATGCGCTGTGATGACCAACACCCCGCGTGGCGGCGACAT 773

```

DB 422 TGCATCTCAACGACGCGACCATCATGACCATTCCTCA-----GACCG 466
QY 774 CTACAAGCGGTGATCATCTGGGCTGAACAAGATGTCGGATGTACAGCCCGTGAG 833
DB 467 CGTAAGCCGTGCGACGCGCCGACTGTGAGAGCTCAAGAGATCTTCCGACGGGAT 526
QY 834 CATCTGACATCCCGCAGGCGCCCAAGAGAGCCCTTCCGCACTACGTGA----- 884
DB 527 CGTGTAGGAGCTTACATGCGCTGCGCACGCGCTCTTCTTACCTGCGCACGACAC 586
QY 885 CCGCTTCTTCAAGACCTGCGCGCGAGCAGCGCACCGAGACGTGAAGAACTGATGAC 944
DB 587 CGACTTCTTCAACGCGCTTCCGCGTGCCTCATCAAGAGAACGACAGAGAGCTGAT 646
QY 945 CGAGACCTGTGTGAGAACGCGCAACCCGCACTGCAAGACATCTGCGGCTCTCGG 1004
DB 647 GCGGCGGTGACCTGCAAGTACATCATGACGCGCTCATCTTGTGACCGGTC 706
QY 1005 CCGCGCGCACCTGAGAGATGATGACCGCTGCGCAGGCGGTGCGCGCGCCCA 1064
DB 707 CCGAGCTGTCTTCTGTGAGCG-----CCGCGCGCGCTGTGTGATCGCTTCT 760
QY 1065 CAAGCGCGCGTGTGCGGAGGATGAGCAGCGCAACAGCGTGAATCATGATGCA 1124
DB 761 GCGGCGCGAGCTGTGTGCGGAGGATGAGCAGCGTGAAGTCTGCAAGAT 820
QY 1125 GAAGAGCACTTCAAGGCGCGCGCGCAAGTCAAGTCTTCACTGCGCAAGAGG 1184
DB 821 GAGGCGCATCGGTGCGGCGTGGGCGCGCGCATCTGGGCGTTCAAGCTTCACTT 880
QY 1185 CCACATGCGCAAGAACTGCGCGCGCGCGCAAGAGGCTGTGAGAGTGCAGCAAGA 1244
DB 881 CCGGTGAGAGTGTCAAGTTGCGCATCGCTGACGCTCTCGGCAAGGCTGGAACA 940
QY 1245 GGGCCACAGATGAAGAGTGAAGCGGCGCGCAAGCTTCTTGGGCAAGATCTGCG 1304
DB 941 CATCAACAACAAGAGCGGCTTCAACGACCGCACTACGCAAGGCGAG----- 992
QY 1305 CAGCCACAAGGCGCGCGCGCACTTCTGCAAGAACGAGCGGCGCGCGCGCAC 1364
DB 993 -----CGAGAGGCGAGTGGCGCACAGAGAGCGTGCACGCGCTCAACAGCGCAC 1048
QY 1365 CGTGCACACGCGCGCGCGCGCGAGCTTCTGAGAGAGCAACCGCGCGCGCA 1424
DB 1049 CGCCACTTCCGACTCTTCTGCGGCAACAGAGGCTTACGCGAGCTGTGAGCTCGCGA 1108
QY 1425 GCAGAGCGCCAGAGAGCGCGAG 1446
DB 1109 GCAGGCGCGCAAGCGCGCGAG 1130

```

RESULT 15  
CG274385 883 bp DNA linear GSS 25-AUG-2003

LOCUS CG274385 883 bp DNA linear GSS 25-AUG-2003  
DEFINITION OG2BH19TV ZM\_0.7\_1.5\_KB Zea mays genomic clone ZMMBMA0752C14,  
genomic survey sequence.

ACCESSION CG274385  
VERSION CG274385.1 GI:34186526

KEYWORDS GSS.  
SOURCE Zea mays  
ORGANISM Zea mays

REFERENCE 1 (bases 1 to 883)  
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Uterback, T.,  
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,  
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.  
Consortium for Maize Genomics

TITLE JOURNAL  
COMMENT Other\_GSSs: OG2BH19TH  
Contact: Cathy Whitelaw

TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whitelaw@tigr.org  
Seq primer: TP  
Class: sheared ends.

FEATURES  
source location/Qualifiers

1..883  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone\_lib="ZMMBMA0752C14"  
/note="Vector: pBSK-; Site\_1: HincII; 0.7-1.5 kb  
methylation filtered genomic DNA library"

# ORIGIN

Query Match 5.0%; Score 75.4; DB 29; Length 883;  
Best Local Similarity 47.0%; Pred. No. 8.8;  
Matches 303; Conservative 0; Mismatches 336; Indels 6; Gaps 2;

```

QY 659 TGGCCCCGCGCAGATGCGCGACCCCGCGGCAAGGACATCGCGCGCCACAGCACCC 718
DB 15 TGGCCAGGCGGAGCTGAGAGACGCGTGAAGAGGCTGCTCGCGCTGAAGCGCGCTTACC 74
QY 719 TGCAGAGCAGATCGCTGTGATGACAGCAACCCCGCTGCGCGCGTGGGCGACATCA 778
DB 75 CGCTGCGCTGCTCCGCGGATGTGAAGAGTTCCCAAGGCGGTGAAGCCCGCGGT 134
QY 779 AGCGGTGATCATCTGCGGCTGAAACAAGATCGTGGGATGTACAGCCCGTGAAGATCC 838
DB 135 GGTGCGCGCGCGCGCGCGCGAGAGAGGCGCAAGAGCGCGAGAGCTCCGCTCAGGA 194
QY 839 TGAACATCGCGCAGGCGCGCGAGAGGCTTCCGCGACTACGTGAAGCGCTTCTTCAAGA 898
DB 195 GCTTGGGCGCGCGCGCGCAATGTGTGGGCGAGAGAGTGAAGAGATGCGCGGCC 254
QY 899 CCTGCGCGCGAGAGGCGCGAGAGCGTGAAGAACTGATGACCGAGACCTGCTG 958
DB 255 TGTGCGCGTGTGCGGCGCA--AGAGAGAGAGAGTACTGAGGTTGGGAAAGCTG 311
QY 959 TGCAGAGCGCAACCGCGACTGCAAGACCATCTGCGCGCTCTCGCGCGCGCGCACCC 1018
DB 312 TGTGAGGCGTCAACGCG--GCTGCGCGTGTGCGGCGCGCGCGCTGAGCGCGCG 368
QY 1019 TGAAGAGATGATACCGCTGCAAGGCGGTGCGCGCGCGCGCGCAAGGCGCGGTG 1078
DB 369 CGGTGCGCACGCGCTTATCGCGGAGAGCGCGCTGCGCGCTCGCGCGCGGTGAGCT 428
QY 1079 TGGCGAGGCGATGAGCGAGCGCAACGCTGAACATCATGATGAGAGAGCAACTTCA 1138
DB 429 GCGGCGCGCTGCGCGCGCGGTCAACAGGTGAGAGAGCGCGCGAGCTGCGCATGCTGT 488
QY 1139 AGGCGCGCGCGCGAGCTCAAGTCTTCACTGCGCGCAAGAGGCGCACATCGCGCAAGA 1198
DB 489 TGAAGTGTGCGCACTGCGCGCGCTTCAACGCAAGCTGCAAGAGAGCATGAGCGCG 548
QY 1199 ACTGCGCGCGCGCGCAAGAGGCTGTGAGAGTGTGCGCGCAAGAGGCGCACAGATGA 1258
DB 549 CGGTGGCGGAGCGCGAGCGTGAAGCGCGGAGAGAGGAGAGCTGTTCGAGCAAGCTG 608
QY 1259 AGGACTGACGAGCGCGAGCGCACTTCTGCGCGAGATCTGCG 1303
DB 609 CGCTGTGCTGTGGCGCGAGCGGTGGAAGCTCGCGCATTCAGGC 653

```

Search completed: June 1, 2004, 14:10:55  
Job time : 3677 secs